

# **Entwicklung und Evaluation eines motorischen Screenings zur Beurteilung der Mobilität bei Erwachsenen**

Dissertation  
zur Erlangung des akademischen Grades eines  
Doktors der Philosophie  
der Fakultät HW  
Bereich Empirische Humanwissenschaften  
der Universität des Saarlandes

vorgelegt von  
Christian Kaczmarek  
aus Neunkirchen/Saar

Saarbrücken, 2018

Dekan: Univ.-Prof. Dr. Cornelius König

Berichterstatter: Univ.-Prof. Dr. Georg Wydra  
Univ.-Prof. Dr. Michael Fröhlich

Tag der Disputation: 23.11.2018

## **Danksagung**

Ich möchte mich bei denjenigen bedanken, die mich während der Promotionszeit begleitet und somit einen großen Teil zum Gelingen dieser Dissertation beigetragen haben.

Mein erster Dank gilt meinem Doktorvater Univ.-Prof. Dr. Georg Wydra, der mir diese Promotion ermöglichte. Vielen Dank für das Vertrauen in mich und die inhaltlichen Diskussionen.

Herrn Univ.-Prof. Dr. Michael Fröhlich danke ich für die Bereitschaft dieses Promotionsverfahren als Zweitgutachter zu übernehmen und für die methodischen Ratschläge.

Herrn Dr. Markus Schwarz bin ich zu besonderem Dank verpflichtet. Herzlichen Dank für die unzähligen Arbeitseinheiten, bei denen neben konstruktiven, wegweisenden Ratschlägen, Korrekturhilfen und Diskussionen auch der Spaß nie zu kurz kam.

Mein weiterer Dank gilt Univ.-Prof. Dr. Eike Emrich für seine methodische und statistische Beratung.

Des Weiteren bedanke ich mich bei meinen aktuellen und ehemaligen Kolleginnen und Kollegen für die schöne Zeit. Christoph Engert danke ich für die praktische Unterstützung bei der Entwicklung des Tests. Dr. Thomas Haab danke ich für die vielen motivierenden Gespräche – es war eine schöne gemeinsame Studienzeit und Promotionsphase.

Des Weiteren bedanke ich mich beim Institut für Sport- und Präventivmedizin der Universität des Saarlandes (Univ.-Prof. Dr. Tim Meyer), der Bosenberg Klinik (Dr. Ulrich Jobst, Sporttherapeuten), der Blietal-Kliniken (Prof. Dr. Stefan Rupp, Sporttherapeuten) und Frau Susanne Rupp. Sie ermöglichten die Durchführung der Untersuchungen in ihren Institutionen und Gesundheits- und Rehabilitationssportgruppen.

Zudem bedanke ich mich bei allen Probanden, die an den Studien teilnahmen und bei den Studierenden, die bei den Datenaufnahmen mithalfen.

Ganz herzlich danke ich meiner Familie, insbesondere meinen Eltern Annemarie und Max, die mir mein Studium ermöglichten, meine Entscheidungen stets respektierten und mich bei meinem Vorhaben unterstützten. Ein besonderer Dank geht auch an meine Freunde, die für die nötige Abwechslung sorgten aber auch Verständnis hatten, wenn die Arbeit im Fokus stand. Alexandra Lieb danke ich für die guten sprachlichen Korrekturen.

Mein größter Dank gilt meiner Freundin Sabine für ihre liebevolle Unterstützung und ihr Verständnis für meine Arbeit während den letzten Jahren.

**Herzlichen Dank für die schöne und lehrreiche Zeit!**

# Inhaltsverzeichnis

<b>VORWORT .....</b>	<b>1</b>
<b>1        EINLEITUNG .....</b>	<b>4</b>
<b>2        MOBILITÄT .....</b>	<b>7</b>
2.1    Einflussfaktoren .....	7
2.1.1    Alter .....	8
2.1.2    Erkrankungen .....	11
2.2    Mobilität und Gesundheit .....	14
2.2.1    Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) .....	14
2.2.2    Mobilität innerhalb der ICF .....	16
2.3    Klassische Motorikkonzepte .....	19
2.3.1    Fähigkeiten und Fertigkeiten .....	19
2.3.2    Kompetenzen .....	22
2.3.3    Aktivitäten des täglichen Lebens .....	24
2.4    Zusammenfassung .....	26
<b>3        SCREENINGTESTS .....</b>	<b>28</b>
3.1    Assessmentprozess .....	28
3.2    Sequentielle Diagnosestrategie .....	30
3.3    Allgemeine Test- vs. Screeningverfahren .....	31
3.3.1    Evaluationskriterien .....	33
3.3.2    Testgüte und diagnostische Güte motorischer Tests und Screenings .....	35
3.3.3    Diagnostische Güte .....	37
3.3.4    Grenzwertproblematik .....	41
3.4    Zusammenfassung .....	46
<b>4        ENTWICKLUNG EINES NEUEN SCREENINGS .....</b>	<b>47</b>
4.1    Klassischer Timed Up and Go Test (TUG) .....	47
4.1.1    Anwendung .....	48
4.1.2    Einflussfaktoren .....	48

4.1.3	Störgrößen .....	50
4.1.4	Testgüte .....	52
4.1.5	Diagnostische Güte .....	54
4.1.6	Testmodifikationen .....	56
4.1.7	Zusammenfassung.....	58
4.2	Erweiterter Timed Up and Go Test (TUG <sub>plus</sub> ) .....	59
4.2.1	Entwicklung ICF-basierter Items .....	59
4.2.2	Ablauf .....	60
4.2.3	Bewertung .....	64
4.3	Zusammenfassung .....	67
5	EVALUATION DES TUG <sub>PLUS</sub> .....	68
5.1	Übergeordnete Fragestellungen und Methodik.....	68
5.1.1	Untersuchungen .....	68
5.1.2	Stichproben.....	69
5.1.3	Verwendete Testverfahren .....	72
5.1.3.1	Fragebögen .....	72
5.1.3.2	Labortest .....	73
5.1.3.3	Feldtests .....	74
5.1.4	Statistik .....	76
5.2	Vor- bzw. Paralleluntersuchung – Klassischer TUG .....	79
5.2.1	Fragestellungen .....	79
5.2.2	Methodik – Personenstichprobe und Untersuchungsablauf .....	80
5.2.3	Hypothesen .....	80
5.2.4	Statistik .....	80
5.2.5	Ergebnisse .....	81
5.3	Evaluation des TUG <sub>plus</sub> anhand verschiedener Alters- und Personengruppen .....	90
5.3.1	Fragestellungen .....	90
5.3.2	Methodik.....	90
5.3.2.1	Personenstichprobe.....	90
5.3.2.2	Variablenstichprobe .....	91
5.3.2.3	Ablauf der Untersuchung .....	91
5.3.3	Hypothesen .....	91

<b>5.3.4</b>	<b>Statistik .....</b>	<b>92</b>
<b>5.3.5</b>	<b>Ergebnisse .....</b>	<b>93</b>
5.3.5.1	<i>Lösungshäufigkeiten .....</i>	93
5.3.5.2	<i>Testzeiten .....</i>	96
5.3.5.3	<i>Testzeiten, Ausführungsökonomie, Extremwert- bzw. Einzelfallbetrachtungen ..</i>	106
<b>5.4</b>	<b>Reliabilität .....</b>	<b>108</b>
<b>5.4.1</b>	<b>Interne Konsistenz .....</b>	<b>108</b>
5.4.1.1	<i>Fragestellungen .....</i>	108
5.4.1.2	<i>Methodik .....</i>	108
5.4.1.2.1	PERSONENSTICHPROBE .....	108
5.4.1.2.2	VARIABLENSTICHPROBE .....	109
5.4.1.2.3	ABLAUF DER UNTERSUCHUNG .....	109
5.4.1.3	<i>Hypothesen .....</i>	110
5.4.1.4	<i>Statistik .....</i>	110
5.4.1.5	<i>Ergebnisse .....</i>	111
5.4.1.5.1	ITEM ZEITEN UND KORRELATIONEN .....	111
5.4.1.5.2	TRENNSCHÄRFE .....	113
<b>5.4.2</b>	<b>Paralleltest-Reliabilität .....</b>	<b>115</b>
5.4.2.1	<i>Fragestellungen .....</i>	115
5.4.2.2	<i>Methodik .....</i>	115
5.4.2.2.1	PERSONENSTICHPROBE .....	115
5.4.2.2.2	VARIABLENSTICHPROBE .....	115
5.4.2.2.3	ABLAUF DER UNTERSUCHUNG .....	115
5.4.2.3	<i>Hypothesen .....</i>	116
5.4.2.4	<i>Statistik .....</i>	117
5.4.2.5	<i>Ergebnisse .....</i>	117
<b>5.5</b>	<b>Responsivität .....</b>	<b>118</b>
<b>5.5.1</b>	<b>Fragestellungen .....</b>	<b>118</b>
5.5.1.1	<i>Methodik .....</i>	118
5.5.1.1.1	PERSONENSTICHPROBE .....	118
5.5.1.1.2	VARIABLENSTICHPROBE .....	118
5.5.1.1.3	TREATMENTSTICHPROBE .....	119
5.5.1.1.4	ABLAUF DER UNTERSUCHUNG .....	120
5.5.1.2	<i>Hypothesen .....</i>	121
5.5.1.3	<i>Statistik .....</i>	121
5.5.1.4	<i>Ergebnisse .....</i>	121
5.5.1.4.1	TESTEFFEKTE .....	121
5.5.1.4.2	RESPONSIVITÄT (NACH ZWEIWÖCHIGER INTERVENTION) .....	122
<b>5.6</b>	<b>Validität .....</b>	<b>123</b>

<b>5.6.1</b>	<b>Kriteriumsvalidität .....</b>	<b>123</b>
5.6.1.1	<i>Fragestellungen.....</i>	123
5.6.1.2	<i>Methodik.....</i>	123
5.6.1.2.1	PERSONENSTICHPROBE .....	123
5.6.1.2.2	VARIABLENSTICHPROBE .....	124
5.6.1.2.3	ABLAUF DER UNTERSUCHUNG .....	124
5.6.1.3	<i>Hypothesen .....</i>	125
5.6.1.4	<i>Statistik .....</i>	125
5.6.1.5	<i>Ergebnisse .....</i>	125
5.6.1.5.1	AUßENKRITERIEN – INTERKORRELATIONEN .....	125
5.6.1.5.2	ZUSAMMENHANG TUG <sub>PLUS</sub> UND AUßENKRITERIEN .....	126
<b>5.6.2</b>	<b>Konstruktvalidität .....</b>	<b>127</b>
5.6.2.1	<i>Fragestellungen.....</i>	127
5.6.2.2	<i>Methodik.....</i>	127
5.6.2.2.1	PERSONENSTICHPROBE .....	127
5.6.2.2.2	VARIABLENSTICHPROBE UND UNTERSUCHUNGSABLAUF .....	128
5.6.2.3	<i>Hypothesen .....</i>	128
5.6.2.4	<i>Statistik .....</i>	128
5.6.2.5	<i>Ergebnisse .....</i>	130
5.6.2.5.1	EXPLORATIVE FAKTORENANALYSE.....	131
5.6.2.5.2	KONFIRMATORISCHE FAKTORENANALYSE .....	134
<b>5.7</b>	<b>Grenzwertwertbestimmung .....</b>	<b>136</b>
<b>5.7.1</b>	<b>Ergebnisverteilung .....</b>	<b>136</b>
5.7.1.1	<i>Fragestellungen.....</i>	136
5.7.1.2	<i>Methodik.....</i>	136
5.7.1.2.1	PERSONENSTICHPROBE .....	136
5.7.1.2.2	VARIABLENSTICHPROBE UND ABLAUF DER UNTERSUCHUNG.....	136
5.7.1.3	<i>Hypothesen .....</i>	136
5.7.1.4	<i>Statistik .....</i>	137
5.7.1.5	<i>Ergebnisse .....</i>	137
<b>5.7.2</b>	<b>Diagnostische Güte .....</b>	<b>138</b>
5.7.2.1	<i>Fragestellungen.....</i>	139
5.7.2.2	<i>Methodik.....</i>	139
5.7.2.2.1	PERSONENSTICHPROBE .....	139
5.7.2.2.2	VARIABLENSTICHPROBE .....	141
5.7.2.2.3	ABLAUF DER UNTERSUCHUNG .....	142
5.7.2.3	<i>Hypothesen .....</i>	143
5.7.2.4	<i>Statistik .....</i>	143
5.7.2.5	<i>Ergebnisse .....</i>	144
5.7.2.5.1	REFERENZSTANDARD „GESUNDHEITSSTATUS“ .....	144

---

5.7.2.5.2	REFERENZSTANDARD „MOTORISCHE LEISTUNGSFÄHIGKEIT“ .....	147
5.7.2.5.3	REFERENZSTANDARD „MOTORISCHE LEISTUNGSFÄHIGKEIT“ UND „GESUNDHEITSSTATUS“ .....	149
5.7.2.5.4	ROC-ANALYSEN .....	150
<b>6</b>	<b>DISKUSSION .....</b>	<b>152</b>
<b>6.1</b>	<b>Evaluation des TUG<sub>plus</sub> .....</b>	<b>154</b>
6.1.1	Timed Up and Go Test .....	154
6.1.2	Timed Up and Go Test plus .....	157
6.1.3	Reliabilität .....	161
6.1.4	Responsivität .....	162
6.1.5	Validität .....	163
6.1.6	Grenzwertbestimmung .....	165
6.1.6.1	<i>Ergebnisverteilung .....</i>	<i>165</i>
6.1.6.2	<i>Diagnostische Güte .....</i>	<i>166</i>
<b>6.2</b>	<b>Methodenkritik .....</b>	<b>171</b>
<b>7</b>	<b>ZUSAMMENFASSUNG UND AUSBLICK .....</b>	<b>174</b>
	<b>LITERATURVERZEICHNIS .....</b>	<b>182</b>



## Abbildungsverzeichnis

Abbildung 1: Gewonnene Lebensjahre in Abhängigkeit von der körperlichen Aktivität .....	8
Abbildung 2: Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit .....	15
Abbildung 3: ICF-Modell mit Hervorhebung der Mobilitätskategorien.....	17
Abbildung 4: Motorische Fähigkeiten .....	20
Abbildung 5: Motorische Fertigkeiten .....	21
Abbildung 6: Bewegungsbezogene Gesundheitskompetenz .....	23
Abbildung 7: Unterkategorien der Mobilität innerhalb der ICF.....	25
Abbildung 8: Diagnostischer Prozess .....	29
Abbildung 9: Sequentielle Diagnosestrategie .....	30
Abbildung 10: Evaluationskriterien für allgemeine Tests und Screeningtests .....	34
Abbildung 11: Vierfeldertafel.....	39
Abbildung 12: Hypothetische Verhältnisannahme von Sensitivität und Spezifität .....	40
Abbildung 13: Grenzwertbestimmung bei normal- und nicht normalverteilten Ergebnissen ....	43
Abbildung 14: Schematischer Aufbau des Timed Up and Go Tests .....	47
Abbildung 15: Weiterentwicklungen des TUG .....	56
Abbildung 16: Mobilitätsdomänen des TUG <sub>plus</sub> .....	61
Abbildung 17: Testaufbau und -ablauf des TUG <sub>plus</sub> .....	62
Abbildung 18: Untersuchungen zur Testevaluation.....	69
Abbildung 19: Messung der Handkraft: .....	75
Abbildung 20: Patientengruppenabhängige Testzeiten beim TUG .....	82
Abbildung 21: Unterschiede der Altersquartile hinsichtlich der Testzeiten beim TUG.....	84
Abbildung 22: Altersabhängige Unterschiede der Testzeiten beim TUG bei „Körperlich Gesunden“ .....	85
Abbildung 23: Altersabhängige Unterschiede der Testzeiten beim TUG bei HNO-Patienten ..	86
Abbildung 24: Altersabhängige Unterschiede der Testzeiten beim TUG bei neurologischen Patienten .....	87
Abbildung 25: Altersabhängige Unterschiede der Testzeiten beim TUG bei Personen mit inneren Erkrankungen .....	88
Abbildung 26: Altersabhängige Unterschiede der Testzeiten beim TUG bei orthopädischen Patienten .....	89
Abbildung 27: Verteilung der Personen- und Patientengruppen in Abhängigkeit der Testzeiten beim TUG <sub>plus</sub> .....	96
Abbildung 28: Testzeiten des TUG <sub>plus</sub> bei unterschiedlichen Personengruppen .....	98
Abbildung 29: Altersspezifische Unterschiede der Testzeiten beim TUG <sub>plus</sub> der Gesamtstichprobe .....	100
Abbildung 30: Neurologie – Altersabhängige Testzeiten beim TUG <sub>plus</sub> .....	101
Abbildung 31: HNO – Altersabhängige Testzeiten beim TUG <sub>plus</sub> .....	102
Abbildung 32: Innere – Altersabhängige Testzeiten beim TUG <sub>plus</sub> .....	103

---

Abbildung 33: Orthopädie – Altersabhängige Testzeiten beim TUG <sub>plus</sub> .....	104
Abbildung 34: „Gesunde“ – Altersabhängige Testzeiten beim TUG <sub>plus</sub> .....	105
Abbildung 35: Zusammenhang zwischen Ausführungsökonomie und Testzeit beim TUG <sub>plus</sub> .....	106
Abbildung 36: Untersuchungsdurchführung mit Itemmessung und -auswertung.....	109
Abbildung 37: Paralleltest-Reliabilität – Untersuchungsdesign, Testablauf .....	116
Abbildung 38: Untersuchungsdesign Testeffekt und Responsivität .....	120
Abbildung 39: Unterschiede der Testzeiten bei den Patientengruppen nach zweiwöchiger Intervention.....	122
Abbildung 40: Scree-Test.....	132
Abbildung 41: Ergebnisse der konfirmatorischen Analyse .....	135
Abbildung 42: Verteilung der Testzeiten der Gesamtstichprobe.....	137
Abbildung 43: Bestimmung von spezifischen Grenzwerten anhand verschiedener Referenz- standards – Gesundheitsstatus und/oder Motorische Leistungsfähigkeit .....	138
Abbildung 44: Testzeiten (TUG <sub>plus</sub> ) von Personen mit und ohne körperlichen Erkrankungen	144
Abbildung 45: Testzeiten des TUG von Personen, die weniger als 21 s oder genau bzw. mehr als 21 s für den TUG <sub>plus</sub> benötigen .....	146
Abbildung 46: Unterschiede der Testzeiten beim TUG <sub>plus</sub> bei Personen, die eine Geschwindigkeit von 7 km/h bei einer stufenförmigen Belastung schaffen und solchen, die dies nicht schaffen .....	148
Abbildung 47: Richtig-, Falsch-Positiv-Raten und ROC-Kurven in Abhängigkeit des Referenzstandards: Berechnete Cut-Off-Werte und Area Under the Curves .	150
Abbildung 48: Richtig-Positiv-Raten des TUG <sub>plus</sub> in Abhängigkeit des Referenzstandards und den ermittelten Cut-Off-Werten .....	151

## Tabellenverzeichnis

Tabelle 1: Häufigkeiten von alters- und geschlechtsspezifischen Einschränkungen bei Alltagstätigkeiten .....	9
Tabelle 2: Prozentualer Anteil an Personen mit chronischen Erkrankungen .....	11
Tabelle 3: Vergleichbare Motorikkonstrukte .....	26
Tabelle 4: Auswahl motorischer Screenings: Studien zur Test- und diagnostischen Güte ....	36
Tabelle 5: Diagnostische Evidenz .....	41
Tabelle 6: Grenzwertbestimmungen .....	42
Tabelle 7: ICF-Grenzwertbestimmung .....	46
Tabelle 8: Altersabhängige Testzeiten beim TUG.....	48
Tabelle 9: Instruktionsabhängige Testzeiten beim TUG.....	51
Tabelle 10: Studien zur Objektivität, Reliabilität und Validität.....	53
Tabelle 11: Studien zur diagnostischen Güte des TUG .....	55
Tabelle 12: Expertenbeurteilung: Vor- und Nachteile der Testaufgaben beim TUG <sub>plus</sub> .....	64
Tabelle 13: Entwicklung eines Beurteilungsbogens .....	65
Tabelle 14: Bewertung der Testaufgaben anhand der Bewegungsqualität .....	66
Tabelle 15: Subjektiver Anstrengungsgrad .....	73
Tabelle 16: Korrelationskoeffizient und entsprechende Interpretation.....	77
Tabelle 17: Signifikanzniveau und entsprechende Interpretation .....	77
Tabelle 18: Effektstärkemaße Eta-Quadrat und Cohens d.....	77
Tabelle 19: Fit Indizes: Formeln und Grenzwerte für die Qualität von Modellanpassungen ....	78
Tabelle 20: TUG: Anthropometrische Daten .....	80
Tabelle 21: Testzeiten des TUG bei verschiedenen Personengruppen .....	81
Tabelle 22: Altersspezifische Testzeiten .....	83
Tabelle 23: Anthropometrische Daten der Gesamtstichprobe beim TUG <sub>plus</sub> .....	90
Tabelle 24: Lösungshäufigkeiten der Items .....	93
Tabelle 25: Lösungshäufigkeiten.....	94
Tabelle 26: Lösungshäufigkeiten der Items .....	95
Tabelle 27: Testindex und Minimum-, Maximumwerte, in Abhängigkeit der Personengruppe, Altersquartile und des Geschlechts.....	95
Tabelle 28: Testzeiten der Teilnehmer, die alle Items absolvieren konnten und geschlechtsabhängige Unterschiede.....	97
Tabelle 29: Altersspezifische Testzeiten von Teilnehmern, die alle Items absolvieren konnten, differenziert nach der Stichprobe .....	99
Tabelle 30: Extremwerte – Patienten mit den langsamsten Testzeiten beim TUG <sub>plus</sub> (≥ 48 s): anthropometrische Daten, Diagnosen, Testzeiten beim klassischen TUG und TUG <sub>plus</sub> inkl. Ausführungsqualität .....	107
Tabelle 31: Interne Konsistenz: Anthropometrische Daten der Personen- bzw. Patientengruppen .....	108
Tabelle 32: Start-, Endpunkte der Messung der Zwischenzeiten bei den einzelnen Items....	110

Tabelle 33: Zwischenzeiten der Items geschlechtsabhängig und -unabhängig.....	111
Tabelle 34: Geschlechtsunabhängige Itemkorrelationen.....	111
Tabelle 35: Itemkorrelationen.....	112
Tabelle 36: Itemkorrelationen.....	112
Tabelle 37: Geschlechtsunabhängige Statistik der jeweiligen Items in Bezug zur Testzusammensetzung .....	113
Tabelle 38: Itemstatistik bei Männern.....	113
Tabelle 39: Itemstatistik bei Frauen .....	114
Tabelle 40: Paralleltest-Reliabilität – Anthropometrische Daten der Patientengruppen.....	115
Tabelle 41: Testzeiten und subjektiver Anstrengungsgrad in Abhängigkeit der Itemreihenfolge. Korrelationen der Variablen zwischen der klassischen und umgekehrten Reihenfolge .....	117
Tabelle 42: Testeffekt und Responsivität – Anthropometrische Daten .....	118
Tabelle 43: Testeffekt: Testzeiten der Patienten beim Vortest 1 und 2 sowie personengruppenabhängige Zusammenhänge.....	121
Tabelle 44: Responsivität: Testzeiten der Patienten.....	122
Tabelle 45: Anthropometrische Daten der Probanden, differenziert nach Validitätskriterien.	123
Tabelle 46: Ergebnisse der Validitätskriterien .....	125
Tabelle 47: Interkorrelationen der Validitätskriterien .....	126
Tabelle 48: Korrelationen zwischen den Testzeiten des TUG <sub>plus</sub> und den Validitätskriterien	126
Tabelle 49: Konstruktvalidität – Anthropometrische Daten der Personengruppen .....	128
Tabelle 50: Kommunalitäten der Items für das Gesamtmodell, „Körperlich Gesunde“ und „Körperlich Kranke“ .....	129
Tabelle 51: Itemkorrelationen der Modelle .....	129
Tabelle 52: Benötigte Testzeiten für die Absolvierung der jeweiligen Items beim Gesamtmodell, bei „Körperlich Gesunden“ und „Körperlich Kranken“ .....	130
Tabelle 53: Faktorladungen .....	133
Tabelle 54: Ergebnisverteilung – Anthropometrische Daten der Gesamtstichprobe .....	136
Tabelle 55: Gesundheitsstatus – Anthropometrische Daten der getesteten Probanden .....	140
Tabelle 56: Motorische Leistungsfähigkeit – Anthropometrische Daten der getesteten Patientengruppen .....	140
Tabelle 57: Motorische Leistungsfähigkeit und Gesundheitsstatus – Anthropometrische Daten getesteten Personen.....	141
Tabelle 58: Diagnostische Güte – Differenzierung von Personen mit und ohne körperliche Beeinträchtigungen bei Cut-Off-Werten zwischen 15 und 30 Sekunden .....	145
Tabelle 59: Prozentualer Anteil gelöster MBD-Items .....	147
Tabelle 60: Diagnostische Güte – Differenzierung von Personen, die eine Geschwindigkeit von 7 km/h bei einer stufenförmigen Beanspruchung schaffen und von solchen, die dies nicht schaffen .....	148
Tabelle 61: Diagnostische Güte – Differenzierung von Personen, die eine Geschwindigkeit von 7 km/h bei einer stufenförmigen Beanspruchung schaffen und körperlich nicht beeinträchtigt sind von solchen, die dies nicht schaffen und gleichzeitig körperlich beeinträchtigt sind .....	149

## Abkürzungsverzeichnis

*	Signifikanz (* = signifikant, ** = sehr signifikant, *** = hoch signifikant)
95-%-CI	95-%-Konfidenzintervall
ADL/ATL	Activities of daily living/Aktivitäten des täglichen Lebens
ACSM	American College of Sports Medicine
$\alpha$	Cronbachs Alpha
AUC	Area Under the Curve
b	Body functions
CFI	Comparative Fit Index
d	Daily activities/Cohens d (Effektstärke)
e	Environmental factors
EFA	Explorative Faktorenanalyse
ERGO	Fahrradergometrie
ETUG	Expanded Timed Up and Go
F	F-Wert (Testprüfgröße der Varianzanalyse)
FE	Fahrradergometrie
ICC	Intraclass Correlation
ICF	International Classification of Functioning, Disability and Health
iTUG	Instrumented Timed Up and Go Test
HF <sub>peak</sub>	Maximal gemessene Herzfrequenz
HNO	Hals-Nasen-Ohrenheilkunde
HK	Handkraft
KFA	Konfirmatorische Faktorenanalyse
$\lambda$	Faktorladung
LR+/-	Positive/negative Likelihood Ratio
M	Mittelwert
MBD	Motorische Basisdiagnostik
MBOR	Medizinisch-beruflich orientierte Rehabilitation
MD	Median
MFI	Mobilitäts-Funktionalitäts-Index
mmHg	Millimeter-Quecksilbersäule (Maßeinheit Blutdruck)
MOBAK	Motorische Basiskompetenzen

---

MOBAQ	Motorische Basisqualifikationen
N	Newton – Maßeinheit für Kraft
n	Stichprobengröße
$\eta^2$	Eta-Quadrat (Effektstärke)
NPV	Negativer Prädiktiver Wert
n.s.	Nicht signifikant
p	Signifikanzwert
$p_i$	Schwierigkeitsindex
P <sub>-langsam</sub>	Personen, die den TUG <sub>plus</sub> in mehr als 21 s absolvierten
PPV	Positiver Prädiktiver Wert
P <sub>-schnell</sub>	Personen, die den TUG <sub>plus</sub> in weniger als 21 s absolvierten
r	Korrelationskoeffizient
RI	Ratz-Index
$r_{it}$	Trennschärfekoeffizient
RMSEA	Root Mean Square Error of Approximation
ROC	Receiver Operating Characteristic Curve
s	Body structures/Sekunden
SD	Standardabweichung
SGB	Sozialgesetzbuch
SEN	Sensitivität
SPE	Spezifität
TLI	Tucker Lewis Index
TUG	Timed Up and Go Test
TUG <sub>plus</sub>	Timed Up and Go Test plus
WHO	Weltgesundheitsorganisation
W/kg KG	Watt pro Kilogramm Körpergewicht
$\chi^2$	Chi-Quadrat-Wert
YI	Youden-Index

## Vorwort

Im Rahmen dieser Arbeit wurde ein motorisches Screening entwickelt und evaluiert. Die theoretischen Grundlagen und die durchgeführten Vor- und Hauptuntersuchungen (inklusive Methodik und Ergebnisse) wurden teilweise auf wissenschaftlichen Kongressen vorgestellt und in Tagungsbänden bzw. Fachzeitschriften veröffentlicht.

In Kapitel 2.2 wird das Konstrukt der Mobilität diskutiert. Dies wurde 2016 im *German Journal of Exercise and Sport Research* und 2017 im Tagungsband *Bewegung, Raum und Gesundheit* veröffentlicht:

- Kaczmarek, C., Schwarz, M., & Wydra, G. (2016a). Mobilitätskonzept in Sporttherapie und -wissenschaft. *German Journal of Exercise and Sport Research*, 46, 268 - 274.
- Wydra, G., & Kaczmarek, C. (2017). Von der Fähigkeits- zur Mobilitätsorientierung in Gesundheits- und Rehasport. In H. Wäsche, G. Sudeck, R. S. Kähler, L. Vogt, & A. Woll (Hrsg.), *Bewegung, Raum und Gesundheit* (S. 111 - 119). Hamburg: Feldhaus.

In Kapitel 3 werden Screeningtests thematisiert. Die Charakteristik dieser Verfahren wurde auf der Jahrestagung *Bewegung, Raum und Gesundheit* der dvs-Kommissionen *Gesundheit und Sport und Raum* vorgestellt und im Abstractband veröffentlicht:

- Wydra, G., & Kaczmarek, C. (2016). Screeningtests in Gesundheits- und Rehasport. In H. Wäsche, G. Sudeck, R. S. Kähler, L. Vogt, & A. Woll (Hrsg.), *Bewegung, Raum und Gesundheit - Wechselwirkungen im Spannungsfeld geänderter Lebensbedingungen und Mobilitäten* (Abstractband, S. 84). Karlsruhe: dvs.

Die Entwicklung der Testaufgaben (Kapitel 4.2.1) wurde 2014 auf dem 23. Deutschen Kongress für Rehabilitationsforschung sowie auf der 17. Jahrestagung der dvs-Kommission *Gesundheit* vorgestellt und im Tagungsband *Arbeit – Gesundheit – Rehabilitation* sowie in der Zeitschrift *Bewegungstherapie und Gesundheitssport* veröffentlicht:

- Kaczmarek, C., Schwarz, M., & Wydra, G. (2014a). Entwicklung eines neuen ICF-basierten Mobilitätstests: Ergebnisse einer ersten Pilotstudie. In Deutsche Rentenversicherung Bund (Hrsg.), *Deutscher Kongress für Rehabilitationsforschung: Arbeit – Gesundheit – Rehabilitation* (S. 152 - 153). Berlin: Deutsche Rentenversicherung Bund.

- Kaczmarek, C., Schwarz, M., & Wydra, G. (2014b). Timed Up and Go Test für Patienten mittleren Alters. *Bewegungstherapie und Gesundheitssport*, 30, 240.

Der Bewertungsmaßstab des Screenings (Kapitel 4.2.3) wurde 2015 auf dem 22. Sportwissenschaftlichen Hochschultag vorgestellt und im Tagungsband *Moving Minds – Crossing Boundaries in Sport Science* publiziert:

- Kaczmarek, C., Fröhlich, M., Schwarz, M., & Wydra, G. (2015). Bewertungsmaßstab eines ICF orientierten Mobilitätstests. In T. Könecke, H. Preuß, & W. I. Schöllhorn (Hrsg.), *Moving Minds - Crossing Boundaries in Sport Science. 22. Sportwissenschaftlicher Hochschultag* (Band 251) (S. 352). Hamburg: Czwalina.

Die Untersuchungen zur Reliabilität (Kap. 5.4), Responsivität (Kap. 5.5), Kriteriumsvalidität (Kap. 5.6.1), Konstruktvalidität (Kap. 5.6.2) und diagnostischen Güte (Kap. 5.7.2) wurden zum Teil 2016 auf der Jahrestagung *Bewegung, Raum und Gesundheit – Wechselwirkungen im Spannungsfeld veränderter Lebensbedingungen und Mobilitäten* der dvs-Kommission Gesundheit und 2017 auf dem 26. Deutschen Kongress für Rehabilitationsforschung sowie auf dem 23. Sportwissenschaftlichen Hochschultag präsentiert und in den Tagungsbänden veröffentlicht:

- Kaczmarek, C., Schwarz, M., Backes, A., & Wydra, G. (2016b). Testgüte und Diagnostische Güte eines neuen Mobilitätsscreenings. In Deutsche Vereinigung für Sportwissenschaft (Hrsg.), *Bewegung, Raum und Gesundheit. Wechselwirkungen im Spannungsfeld geänderter Lebensbedingungen und Mobilitäten. Jahrestagung der dvs-Kommissionen „Gesundheit“ und „Sport und Raum“* (S. 57). Karlsruhe: dvs.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2017a). Reliabilität, Validität und diagnostische Güte eines neuen Mobilitätsscreenings. In H. Wäsche, G. Sudeck, R. S. Kähler, L. Vogt, & A. Woll (Hrsg.), *Bewegung, Raum und Gesundheit* (S. 143 - 152). Hamburg: Feldhaus.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2017b). TUG 50+ – Ein neues Screening für die Bewegungstherapie. In A. Schwirtz, F. Mess, Y. Demetriou, & V. Senner (Hrsg.), *Innovation & Technologie im Sport* (S. 274). Hamburg: Feldhaus.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2017c). Teststatistische Überprüfung eines neuen ICF orientierten, modifizierten Timed Up and Go Tests. In Deutsche Rentenversicherung Bund (Hrsg.), *Deutscher Kongress für Rehabilitationsforschung*.



*Prävention und Rehabilitation in Zeiten der Globalisierung*  
(S. 71 - 72). Frankfurt: DRV-Schriften.

In den entsprechenden Kapiteln wird auf die aufgeführten Publikationen hingewiesen.

# 1 Einleitung

Im Jahr 2013 betrug der Bevölkerungsanteil von Personen über 60 Jahren 27 Prozent. Im Zuge des demographischen Wandels wird für diese Altersgruppe bis 2030 ein Anstieg von acht Prozent prognostiziert. Für Personen über 80 Jahre (4,4 Millionen, Stand 2013) wird sogar ein Anstieg um bis zu 40 Prozent erwartet (Statistisches Bundesamt, 2015, S. 19, 45). Die steigende Lebenserwartung und sinkende Geburtenrate führen zu einem immer höheren Altersdurchschnitt der Bevölkerung. Diese Entwicklung bedeutet, dass die Krankheitsprävalenz zunimmt und die Leistungsfähigkeit, aber auch Mobilität altersentsprechend abnehmen. Das Gesundheitssystem wird daher vor neue Herausforderungen gestellt, beispielsweise durch eine erhöhte Inanspruchnahme gesundheitsorientierter Leistungen (Robert Koch Institut, 2015). Hierbei ist die körperliche Leistungsfähigkeit im Gegensatz zu Alterserscheinungen, wie beispielsweise geringe Sehkraft und schlechtes Hörvermögen, recht gut trainierbar. Folglich kann durch gezielte körperliche Aktivität möglichen Einschränkungen entgegengewirkt werden (Garber et al., 2011; Verdijk, Snijders, Holloway, Van, & Van, 2016). Über die Lebensspanne hinweg mobil zu bleiben, stellt somit eine gesellschaftliche Ressource dar.

Motorik bzw. Mobilität entwickeln sich im Kindesalter, werden im Laufe der Zeit automatisiert und nehmen im Alter wieder qualitativ ab (Hardy, Dubin, Holford, & Gill, 2005; Stuck, Walthert, Nikolaus, Büla, Hohmann, & Beck, 1999). Nicht mehr bzw. weniger mobil zu sein, kann dann zu einer zunehmenden Abhängigkeit führen und die Lebensqualität beeinträchtigen (Hardy et al., 2005; Means, Rodell, & O'Sullivan, 2005). Dass Mobilitätseinschränkungen nicht ausschließlich im höheren Alter auftreten, zeigt eine Studie von Iezzoni, McCarthy, Davis und Siebens (2001). So berichten bereits Personen jünger als 60 Jahre über Beeinträchtigungen der Mobilität.

Zur Diagnose können sogenannte motorische Screeningtests einen ersten Überblick über mögliche körperliche Einschränkungen geben. Aufgrund des geringen Material- und Zeitbedarfs ermöglichen solche Verfahren die motorische, körperliche Leistungsfähigkeit bzw. Mobilität auch bei relativ großen Stichproben ökonomisch zu überprüfen (Croyle, 1995, S. 3). Um umfassendere Ergebnisse beispielsweise zur Ausdauerleistungsfähigkeit zu erhalten, sind vor allem ergometrische Stufentests auf dem Laufband oder Fahrrad, bei denen die maximale Herzfrequenz ein Indikator für die Ausbelastung darstellt, etablierte Ver-

fahren (Such, & Meyer, 2010). Durch den in der Internationalen Klassifikation für Funktionsfähigkeit, Behinderung und Gesundheit (ICF – International Classification of Functioning, Disability and Health) vordefinierten Begriff der Mobilität, d. h. die Fähigkeit, über Bewegung mit der Umwelt zu interagieren, der als zentrales Kriterium von Gesundheit durch die Weltgesundheitsorganisation (WHO) international verankert wurde (WHO, 2001), ergeben sich eine Reihe von Anknüpfungspunkten, vor allem für den Gesundheitsbereich.

Für dieses Setting sind in erster Linie ökonomische Screeningverfahren geeignet. Bei Hochaltrigen der Geriatrie ist der Timed Up and Go Test (TUG) von Podsiadlo und Richardson (1991) ein international häufig eingesetztes motorisches Screening zur Beurteilung der Mobilität und des Sturzrisikos. Anhand eines Grenzwertes (Cut-Off-Wert) können hierbei Testergebnisse als auffällig oder nicht auffällig charakterisiert werden. Da der TUG hauptsächlich in der Geriatrie selektiv ist, fehlt es an geeigneten, ökonomischen Methoden zur Messung der Mobilität von Erwachsenen mittleren Alters. Daher soll in der vorliegenden Arbeit ein ICF-orientierter motorischer Screeningtest zur Beurteilung der Mobilität von Personen mittleren Alters entwickelt und evaluiert werden. Der TUG dient hierbei als Basistest und wird zum Timed Up and Go Test plus (TUG<sub>plus</sub>) modifiziert, indem mehr Bewegungen mit Bezug zur Alltagsmobilität integriert werden. In Anlehnung an den TUG soll auch für den TUG<sub>plus</sub> ein Cut-Off-Wert bestimmt werden, anhand dessen motorische Einschränkungen (Mobilität) erkannt werden können.

Die Arbeit gliedert sich wie folgt:

Um die Entwicklung des Screenings thematisch einzuordnen, wird zunächst die Mobilität als theoretische Basis thematisiert (Kap. 2). Hierbei erfolgt eine Auflistung von Faktoren, die die Mobilität beeinflussen können (Kap. 2.1). Diese werden insbesondere hinsichtlich Alter (Kap. 2.1.1) und Erkrankungen (Kap. 2.1.2) diskutiert. Anschließend wird die Mobilität begrifflich und inhaltlich einer Klassifikation der WHO im Gesundheitssektor zugeordnet (Kap. 2.2). In Kapitel 2.3 wird ein Vergleich zu klassischen Motorikkonzepten gezogen. Dabei sind motorische Fähigkeiten und Fertigkeiten, Kompetenzen und Aktivitäten des täglichen Lebens Diskussionsgegenstand.

Anschließend werden Screeningtests thematisiert (Kap. 3). Zunächst erfolgt eine Beschreibung von Assessmentprozessen (Kap. 3.1) und eine Einordnung von Screenings in sequentielle Diagnosestrategien (Kap. 3.2). Anschließend werden Screeningtests inhaltlich und begriff-

lich von klassischen Tests abgegrenzt und die Besonderheiten aufgezeigt (Kap. 3.3). Dies geschieht insbesondere hinsichtlich der Evaluation (Kap. 3.3.1). Daran schließt sich eine Auflistung motorischer Screeningverfahren und deren Test- und diagnostischer Güte an (Kap. 3.3.2). Weiterhin werden die Überprüfung der diagnostischen Güte und die Bestimmung von Grenzwerten (Cut-Off-Werte) beschrieben und im Rahmen der Grenzwertproblematik diskutiert (Kap. 3.3.3; 3.3.4).

Im vierten Kapitel werden die Schritte zur Entwicklung des Screenings (TUG<sub>plus</sub>) aufgezeigt. Dabei erfolgt zunächst die Beschreibung des klassischen TUG (Kap. 4.1) und dessen praktische Anwendung, Einflussfaktoren, Störgrößen sowie Studien zur Testgüte, diagnostischen Güte und bisherigen Testmodifikationen (Kap. 4.1.1 bis 4.1.7). Auf dieser Basis werden in Kapitel 4.2 die Testaufgaben des TUG<sub>plus</sub> vorgestellt und durchgeführte Vorstudien aufgezeigt.

In Kapitel 5 werden nach einer Beschreibung der übergeordneten Fragestellungen und Methodik (Kap. 5.1) die Untersuchungen zur Evaluation des neuen Screenings vorgestellt: Vor- bzw. Paralleluntersuchung zum klassischen TUG (Kap. 5.2), Evaluation des TUG<sub>plus</sub> anhand verschiedener Alters- und Personengruppen (Kap. 5.3), Reliabilität (Kap. 5.4), Responsivität (Kap. 5.5), Validität (Kap. 5.6) und Grenzwertbestimmungen (Kap. 5.7).

Abschließend werden die Ergebnisse diskutiert (Kap. 6) und die wesentlichen Befunde zusammengefasst (Kap. 7), die abschließend in der praktischen Anwendung – einem TUG<sub>plus</sub> Testmanual und einem -protokoll – münden.

## 2 Mobilität

Mit „Mobilität“ werden je nach Zusammenhang unterschiedliche Bedeutungen assoziiert. So finden sich in sozialwissenschaftlichen Kontexten die Begriffe soziale und räumliche Mobilität (Clark, van Ham, & Coulter, 2014; Musterd, Van Gent, Das, & Latten, 2016). Soziale Mobilität bezeichnet die Veränderung der sozialen Stellung in einer Gesellschaft. Räumliche Mobilität impliziert eine örtliche bzw. räumliche Veränderung.

In der Sportwissenschaft wurde der Begriff in der Vergangenheit häufig im Zusammenhang mit Beweglichkeit und Dehntraining verwendet (Godges, MacRae, Longdon, Tinberg, & MacRae, 1989; Hartley-O'Brien, 1980; Williams, & Goldspink, 1978). Im Bereich Motorik definieren Shumway-Cook und Woollacott (2007, S. 300) Mobilität als Fähigkeit, sich sicher und unabhängig von einem Standort zu einem anderen zu bewegen. Der Mobilität ordnen sie die Fertigkeiten Gehen, Laufen, Aufstehen vom Boden/Stuhl/Bett und Rollen zu. Auch die Orientierung im Raum bzw. Wahrnehmungsfähigkeit fassen sie unter dem Begriff zusammen.

Eine genauere Differenzierung des Begriffes beschreibt die Weltgesundheitsorganisation (WHO, 2001) in ihrem Klassifikationssystem der ICF: „Bewegung durch Änderung der Körperposition oder -lage oder Verlagerung von einem Platz zu einem anderen, mit der Bewegung von Gegenständen durch Tragen, Bewegen oder Handhaben, mit der Fortbewegung durch Gehen, Rennen, Klettern oder Steigen sowie durch den Gebrauch verschiedener Transportmittel“ (DIMDI, 2005, S. 103). Dieses Begriffsverständnis kommt in der vorliegenden Arbeit zur Anwendung und wird im Folgenden näher erläutert.

Zunächst werden Faktoren beschrieben, die die Mobilität eines Menschen negativ beeinflussen können. Anschließend erfolgt die inhaltliche Einordnung in die ICF. Die klassischen Motorikmodelle (motorische Fähigkeiten, Fertigkeiten, Kompetenzen, Aktivitäten des täglichen Lebens) werden abschließend dargestellt, um den in dieser Arbeit verwendeten Ansatz einzugrenzen bzw. Gemeinsamkeiten herauszuarbeiten.

### 2.1 Einflussfaktoren

Mobil zu sein befähigt dazu Alltagstätigkeiten verrichten zu können (Hirvensalo, Rantanen, & Heikkinen, 2000, S. 493). Jede Einschrän-

kung der Mobilität kann deren Ausübung behindern. Das Zusammenspiel zwischen Muskulatur und zentralem Nervensystem ermöglicht die Ausführung motorischer Handlungen und kann durch verschiedene Faktoren limitiert werden.

### 2.1.1 Alter

Moore et al. (2012) untersuchten, ob ein zusätzlicher Energieumsatz durch körperliche Aktivität die Lebenserwartung beeinflusst. Sie testeten hierzu rund 655000 Personen im Alter zwischen 21 und 90 Jahren. Die körperliche Aktivität wurde durch das metabolische Äquivalent (englisch: Metabolic Equivalent; MET) angegeben. Ein MET entspricht dem Energieverbrauch beim Sitzen (1 kcal/kg KG/Stunde; 3,5 ml O<sup>2</sup>/kg KG/Min; Ainsworth, Haskell, Leon, Jacobs, Montoye, Sallis, & Paffenbarger, 1993, S. 72). Moore et al. (2012) konnten zeigen, dass ein wöchentlicher Kalorienmehrverbrauch von etwa 2000 kcal die Lebenserwartung um bis zu fünf Jahre erhöhen kann (Abb. 1).

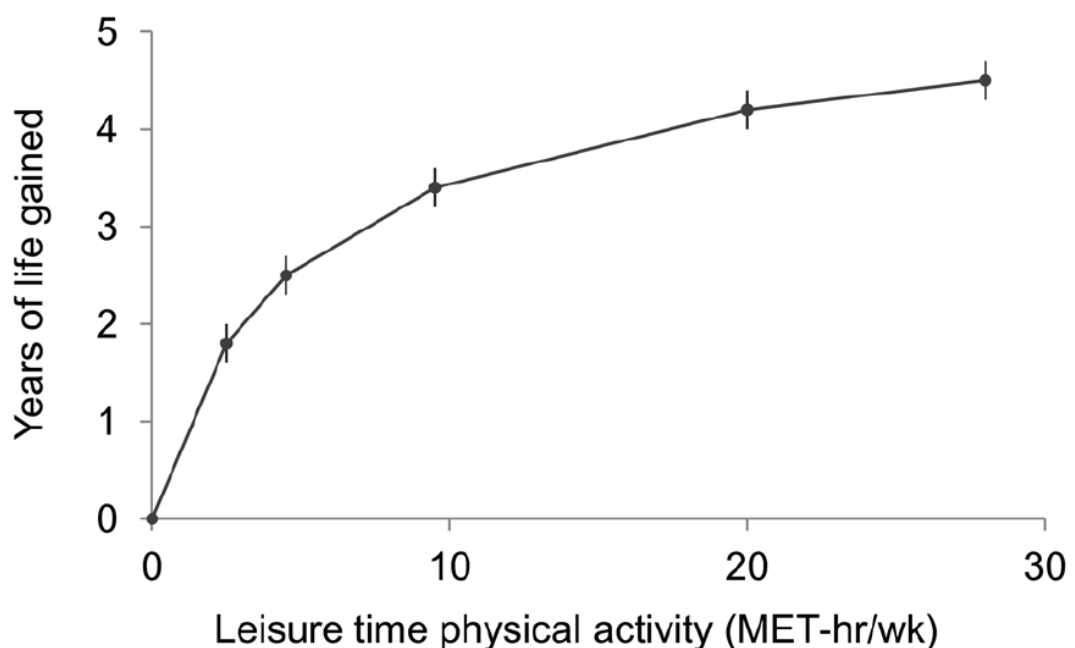


Abbildung 1: Gewonnene Lebensjahre in Abhängigkeit von der körperlichen Aktivität pro Woche (Moore et al., 2012; S. 6)

Es ist somit Voraussetzung im Alter mobil zu bleiben, um den Energieverbrauch auch dann anzukurbeln und einen Gewinn an Lebensjahren zu ermöglichen. Jedoch berichten etwa ein Drittel der 65-Jährigen über Beeinträchtigungen, beispielsweise beim Treppensteigen und Gehen (Shumway-Cook, Ciol, Yorkston, Hoffman, & Chan, 2005). Funktionale Einschränkungen für die Altersspanne 40 bis 85 Jahre wurden bereits im Deutschen Alterssurvey 2008 veröffentlicht (Tab. 1).

*Tabelle 1: Häufigkeiten (Prozent) von alters- und geschlechtsspezifischen (♂, ♀) Einschränkungen bei Alltagstätigkeiten (Selbstausskunft; modifiziert nach Motel-Klingebiel, Wurm, & Tesch-Römer, 2010, S. 21, Deutscher Alterssurvey)*

Tätigkeit	Einschränkung	Altersgruppe (Jahre)					
		40 - 54		55 - 69		70 - 85	
		♂	♀	♂	♀	♂	♀
Anstrengende Tätigkeit	Etwas eingeschränkt	25	35	34	42	46	42
	Stark eingeschränkt	11	13	22	22	35	46
Einkaufstaschen Heben/tragen	Etwas eingeschränkt	5	11	10	20	20	38
	Stark eingeschränkt	1	4	4	6	8	18
Einen Treppenabsatz steigen	Etwas eingeschränkt	3	4	6	9	15	23
	Stark eingeschränkt	1	1	3	3	7	9

Es zeigt sich, dass mit zunehmendem Alter die Einschränkung bei den aufgelisteten Tätigkeiten prozentual zunimmt. Zudem ist zu erkennen, dass Frauen größere Schwierigkeiten bei diesen Tätigkeiten haben als Männer. Mögliche Ursachen werden im Folgenden beschrieben.

### *Kraftniveau*

Frontera, Hughes, Lutz, und Evans (1991) untersuchten die Muskelkraft von 200 gesunden Personen zwischen 45 und 78 Jahren. Es zeigte sich ein mit höherem Alter zunehmender Kraftverlust in fast allen Muskelgruppen. Weiterhin stellten Warburton, Gledhill und Quinney (2001) in einem systematischen Review eine Abnahme der muskulären Kraft von 3,5 % pro Jahr zwischen dem 65. und 84. Lebensjahr fest.

Zur Messung der körperlichen Kraft ist die Handkraft im Gegensatz zu umfangreichen Tests eine ökonomische Alternative, da diese als Indikator für das Kraftniveau des ganzen Körpers gilt (Sasaki, Kasagi, Yama-

da, & Fujitsu, 2007, S. 337). Zudem zeigt sich sogar ein Zusammenhang zur Mortalitätsrate (Rantanen, 2003; Ling, Taekema, de Craen, Gussekloo, Westendorp, & Maier, 2010). Anhand von 2600 gesunden Personen älter als 55 Jahre, wurde ähnlich wie beim TUG (Kap. 1, S. 5) für die Handkraft ein Cut-Off-Wert bestimmt, um Personen mit Mobilitätseinschränkungen (Beeinträchtigungen beim Gehen oder Treppensteigen) identifizieren zu können. Es zeigte sich, dass bei Männern mit einem Wert unter 37 kg und bei Frauen unter 21 kg eine erhöhte Wahrscheinlichkeit für Mobilitätseinschränkungen bestand (Sallinen, Stenholm, Rantanen, Heliövaara, Sainio, & Koskinen, 2010).

### *Koordination – Gleichgewichtsfähigkeit*

Die Fähigkeit „Gehen können“ ist eine wesentliche Voraussetzung, um am sozialen Leben teilhaben zu können. Gehen mit einer geringen ausgeprägten Gleichgewichtsfähigkeit und einer damit verbundenen Sturzangst limitiert die Durchführung alltäglicher Bewegungshandlungen. Dies war das Ergebnis einer Studie, in der die Sturzangst und deren Auswirkungen auf die Mobilität untersucht wurden. Cumming, Salkeld, Thomas und Szonyi (2000) konnten diese Erkenntnisse bestätigen. In ihrer Untersuchung verglichen sie zehn Aktivitäten des täglichen Lebens (ATL; englisch: activities of daily living – ADL) mit dem Tinetti Falls Efficacy Scale (FES), ein Test zur Überprüfung der Gleichgewichtsfähigkeit (Tinetti, Richman, & Powell, 1990).

Die motorischen Einschränkungen stehen in Wechselwirkung mit Erkrankungen. So erhöht sich mit zunehmendem Alter die Prävalenz (Robert Koch Institut, 2015). Im Deutschen Alterssurvey aus dem Jahr 2014 finden sich Angaben zur Anzahl der Erkrankungen von Personen im Alter zwischen 40 und 85 Jahren (Deutsches Zentrum für Altersfragen, 2014, S. 23). Innerhalb dieser Altersspanne litten 38 % an keiner oder einer, 49 % an zwei bis vier und 13 % an mindestens fünf chronischen Erkrankungen (Tab 2).



*Tabelle 2: Prozentualer Anteil an Personen mit chronischen Erkrankungen, differenziert nach Altersspannen (modifiziert nach Deutsches Zentrum für Altersfragen, 2014, S. 23)*

Alter (Jahre)	Anzahl chronischer Erkrankungen	Anteil an Personen mit chronischen Erkrankungen (%)
40 - 54	0 - 1	53
	2 - 4	42
	≥ 5	5
55 - 69	0 - 1	34
	2 - 4	51
	≥ 5	14
70 - 85	0 - 1	18
	2 - 4	57
	≥ 5	25

Im Folgenden werden Erkrankungen und mögliche Auswirkungen auf die Motorik, insbesondere Mobilität, beschrieben.<sup>1</sup>

## 2.1.2 Erkrankungen

### *Innere Erkrankungen*

Innere Erkrankungen, insbesondere die koronare Herzkrankheit, zählen zu den häufigsten Todesursachen in Deutschland (Stand 2013, RKI, 2015, S. 24). So starben 40 % an Erkrankungen des Herz-Kreislaufsystems und 25 % an den Folgen einer Krebserkrankung. Die koronare Herzkrankheit (KHK) bildet bei beiden Geschlechtern (Männer 16 %, Frauen 13 %) die häufigste Todesursache. Bei Männern ist Lungenkrebs die zweithäufigste (7 %) und bei Frauen Brustkrebs die sechsthäufigste Ursache (3 %).

Patienten mit solchen Erkrankungen können eine geringere gewichtsbezogene Leistungsfähigkeit (Watt pro Kilogramm Körpergewicht – W/kg KG; < 1 W/kg KG) als „gesunde“ Personen (Frauen: 2,5 W/kg KG, Männer: 3 W/kg KG) aufweisen (Löllgen, 2010, S. 69) und somit in ihren Alltagstätigkeiten eingeschränkt sein (Graf, & Halle, 2007, S. 324). So haben Personen mit koronaren Herzkrankheiten und Diabetes mellitus ein erhöhtes Risiko für motorische Einschränkungen (Dunlop, Manheim, Sohn, Liu, & Chang, 2002).

Auch Frauen mit Brustkrebs zeigen aufgrund körperlicher Schmerzen vermehrt Mobilitätsprobleme (Kroenke, Rosner, Chen, Kawachi, Colditz,

<sup>1</sup> Die Auswahl beschränkt sich auf Erkrankungen, die in den empirischen Untersuchungen dieser Arbeit thematisiert werden.

& Holmes, 2004). Bei diesem Krankheitsbild führt zudem häufig das sogenannte Fatigue-Syndrom zu einer geringeren körperlichen Leistungsfähigkeit (Dimeo, Stieglitz, Novelli-Fischer, Fetscher, Mertelsmann, & Keul, 1997).

### *Neurologische Erkrankungen*

Schlaganfall ist neben der koronaren Herzkrankheit weltweit die häufigste Todesursache (Lozano et al., 2012, S. 2112). So liegt die Sterberate bei bis zu einem Drittel. Personen, die einen Anfall überleben, leiden häufig an teils schwerwiegenden körperlichen und motorischen Beeinträchtigungen (Kolominsky-Rabas et al., 1998; Palm et al., 2010). Im Gegensatz zu Herzpatienten, bei denen insbesondere die körperliche Leistungsfähigkeit (W/kg KG) erniedrigt ist, liegt bei Schlaganfallpatienten neben dieser Einschränkung häufig auch eine beeinträchtigte Motorik vor. Dies zeigt auch eine Studie zur Alltagsmobilität von 79 Schlaganfallpatienten (Wendel, Stähl, Risberg, Pessah-Rasmussen, & Iwarsson, 2010). 44 % der Personen berichten über Probleme beim Gehen und weisen ein erhöhtes Sturzrisiko auf. Als Ursachen können Einschränkungen der ADL (hauptsächlich beim Gehen), Gleichgewichtsprobleme und Schwindel genannt werden (Harris, Eng, Marigold, Tokuno, & Louis, 2005; Hyndman, Ashburn, & Stack, 2002; Lamb, Ferrucci, Volapto, Fried, & Guralnik, 2003).

Zur Bestimmung des Mobilitätsstatus bei Patienten mit multipler Sklerose dienen die Sturzhäufigkeit und deren Ursachen als Indikator. Es zeigt sich, dass Patienten mit einem erhöhten Sturzrisiko auch vermehrt unter Mobilitätsproblemen leiden (Benutzung von Gehhilfen, geringere Gehausdauer, Gleichgewichtsprobleme; Matsuda, Shumway-Cook, Ciol, Bombardier, & Kartin, 2012; Sosnoff et al., 2011).

In diesem Zusammenhang wird bereits hier darauf hingewiesen, dass in der vorliegenden Arbeit hauptsächlich gehfähige Patienten in die empirischen Untersuchungen mit aufgenommen wurden.

### *Hals-Nasen-Ohren Erkrankungen*

Erkrankungen dieser Bereiche sind sehr vielschichtig. Beispielsweise werden Personen mit Tinnitus, Hörschaden und Schwindel unter dieser Kategorie zusammengefasst. Oft treten auch mehrere Einschränkungen gleichzeitig auf. So leiden Personen mit Morbus Menière häufig an Tinnitus (Wahrnehmung eines Geräusches ohne externe Ursache), Schwindel und Hörverlust (Schaaf, 2017, S. 43 ff.). Studien zufolge sind Personen mit Schwindel vermehrt sturzgefährdet (Herdman, Blatt,

Schubert, & Tusa, 2000; Whitney, Hudak, & Marchetti, 2000) und zeigen Probleme bei der Ausübung alltäglicher Aktivitäten (Hinlegen auf Bett, Aufstehen von Bett/Stuhl, Gehen kleiner Strecken, Stehen über längere Zeit, Bücken und Knien; Harun, Semenov, & Agrawal, 2015).

### *Orthopädische Erkrankungen und Verletzungen*

Arthrose zählt weltweit zu den häufigsten Erkrankungen der Gelenke (WHO, 2003, S. 1). Nach Angaben des Robert Koch-Instituts litten im Jahr 2015 knapp 20 % der Erwachsenen in Deutschland an arthrotischen Erkrankungen (GEDA-Studie). Ab einem Alter von 65 Jahren liegt die Prävalenz bei Frauen bei knapp 50 % und bei Männern bei 30 % (Fuchs, Kuhnert, & Scheidt-Nave, 2017, S. 56 f.). Dabei sind vor allem Knie- (Gonarthrose) und Hüftgelenke (Coxarthrose) betroffen.

Im fortgeschrittenen Stadium können infolge von Schmerzen motorische Einschränkungen entstehen (RKI, 2013, S. 9), die aus medizinischer Sicht oftmals nur durch endoprothetische Versorgung zu beheben sind (Merx, Dreinhöfer, & Günther, 2007, S. 423). Hierbei zeigt sich, dass Personen mit einem präoperativ besseren motorisch, körperlichen Ausgangsniveau nach Einsatz einer Totalendoprothese (Knie oder Hüfte) von weniger Schmerzen, einer höheren Lebensqualität und geringeren motorischen Einschränkungen berichten (Fortin et al., 1999).

Neben Erkrankungen führt die WHO in einem Klassifikationssystem weitere Bestandteile auf, die für ein unabhängiges Leben von Bedeutung sind. Anhand dieser Klassifikation können Gesundheitsprobleme eines Menschen international vergleichbar beschrieben werden.

## 2.2 Mobilität und Gesundheit<sup>2</sup>

### 2.2.1 Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF)

Die WHO konzipierte im Jahr 1948 die Internationale Klassifikation der Krankheiten und verwandter Gesundheitsprobleme (ICD). Die ICD geht auf das von Jacques Bertillon konzipierte „Verzeichnis der Todesursachen“ zurück. Dadurch ermöglichte die WHO neben einer einheitlichen Klassifizierung von Todesursachen auch die von Krankheiten. So können heute medizinische Diagnosen klassifiziert und Krankheiten über ein Verschlüsselungssystem eindeutig beschrieben werden (Osholt-Corsten, & Schuntermann 2011, S. 66 f.).

Später veröffentlichte die WHO die Internationale Klassifikation der Schädigungen, Fähigkeitsstörungen und Beeinträchtigungen (ICIDH: International Classification of Impairments, Disabilities and Handicaps). Neben der Beschreibung von Krankheiten sollten hierdurch insbesondere die Krankheitsfolgen Berücksichtigung finden (DIMDI, 2005).

Nachdem die ICIDH wegen ihrer einfachen Sichtweise bezüglich Störungen infolge von Krankheiten und der fehlenden Berücksichtigung von umwelt- und personenbezogenen Faktoren in der Kritik stand, wurde diese zur ICIDH-2 modifiziert (DIMDI, 2005).

Im Jahr 2001 wurde die ICIDH-2 zur ICF umbenannt. Die ICF sollte dabei fortan zusammen mit der ICD international für ein sprachliches Verständnis sorgen. Die neue Klassifikation sollte eine mehrperspektivische Herangehensweise an Gesundheit und Behinderung ermöglichen. Dominiert in der ICD ein medizinisches Krankheitsparadigma, sind in der ICF neben medizinischen Aspekten auch andere für das menschliche Dasein wichtige Bereiche herausgearbeitet. Auf der Basis eines ganzheitlichen Gesundheitsverständnisses sollte durch diese Klassifikation ermöglicht werden, Auswirkungen von Gesundheitsproblemen auf Komponenten der Funktionsfähigkeit und Behinderung zu erkennen und diese vor dem Hintergrund personen- und umweltbezogener Kontextfaktoren zu betrachten (Stucki, Cieza, & Melvin, 2007, S. 279 f.).

Abbildung 2 zeigt die Bereiche, die aufgrund einer Gesundheitsstörung betroffen sein können oder diese begünstigen.

---

<sup>2</sup> Das Kapitel wurde teilweise im German Journal of Exercise and Sport Research (Kaczmarek et al., 2016a) und im Tagungsband *Bewegung, Raum und Gesundheit* (Wydra, & Kaczmarek, 2017) publiziert.

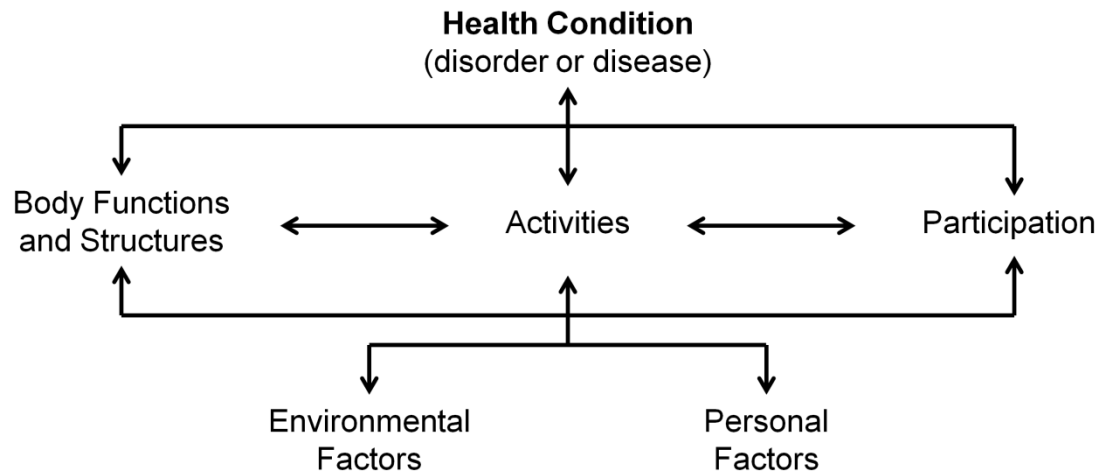


Abbildung 2: Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF; WHO, 2001, S. 18)

Komponenten der Körperfunktionen und -strukturen (DIMDI, 2005, S. 17 ff.):

- *Body functions*: Körperfunktionen – Physiologische und psychologische Funktionen des Körpers (z. B. Seh- und Sprechvermögen, Atmung).
- *Body Structures*: Körperstrukturen – Anatomische Aspekte (z. B. Knochen, Herz).

Komponenten der Aktivitäten und Partizipation:

- *Activities*: Aktivitäten – Durchführung bestimmter Tätigkeiten (z. B. Schwimmen, Einkaufen).
- *Participation*: Partizipation – Teilhabe am gesellschaftlichen Leben (z. B. Alltag, Beruf, Sport).

Umwelt- und personenbezogene Faktoren

- *Environmental Factors*: Umweltfaktoren – „Materielle, soziale und einstellungsbezogene“ Aspekte der Lebensumwelt eines Menschen (DIMDI, 2005, S. 21). Diese lassen sich untergliedern in individuums- (z. B. Familie, Arbeitsplatz) und gesellschaftsbezogene Faktoren (z. B. soziale Netzwerke, Gesetze).
- *Personal Factors*: Personenbezogene Faktoren – z. B. Alter, Geschlecht, unabhängig von Gesundheitsproblemen (DIMDI, 2005, S. 22). Diese können zusammen mit Umweltfaktoren die Funktionsfähigkeit negativ (Barrieren) oder positiv (Förderfaktoren) beeinflussen (Grotkamp et al., 2014, S. 275). Im Gegensatz zu den anderen ICF-Bereichen wurden die personenbezogenen Faktoren nicht klassifiziert. Sie werden jedoch primär im klini-

schen Kontext bei der Beurteilung eines Gesundheitsproblems, der Formulierung von individuellen Zielen und Interventionsinhalten berücksichtigt.

Nach der WHO gilt eine Person als funktional gesund (DIMDI, 2005, S. 4), wenn auf den Ebenen der Körperfunktionen, -strukturen, Aktivitäten und Partizipation keine Beeinträchtigungen vorliegen. Bestehen Einschränkungen auf den Ebenen der Körperfunktionen und/oder -strukturen, müssen diese nicht zwingend auf Aktivitäts- und Partizipationsniveau nachweisbar sein (Kaczmarek, Schwarz, & Wydra, 2016a, S. 269).

Unter welchen Komponenten die Mobilität einzuordnen ist, wird im Folgenden dargestellt.

### **2.2.2 Mobilität innerhalb der ICF**

Mobilität im Sinne der ICF bedeutet, die eigene „Bewegung durch Änderung der Körperposition oder -lage oder Verlagerung von einem Platz zu einem anderen, mit der Bewegung von Gegenständen durch Tragen, Bewegen oder Handhaben, mit der Fortbewegung durch Gehen, Rennen, Klettern oder Steigen sowie durch den Gebrauch verschiedener Transportmittel“ (DIMDI, 2005, S. 103).

Die Mobilität ist den ICF-Komponenten Aktivitäten und Partizipation zugeordnet (Abb. 2, S. 15) und bildet darin eine eigenständige Kategorie neben der Kommunikation, Selbstversorgung, interpersonellen Interaktion und Gemeinschaftsleben (Stucki et al., 2007, S. 280). Obwohl Aktivitäten und Partizipation innerhalb der ICF als eigenständige Komponenten aufgeführt sind, werden diese in einer gemeinsamen Komponente (daily activities – d) aufgeführt.

Aufgrund unterschiedlicher Definitionen stellt sich die Frage, worin sich Aktivitäten und Partizipation unterscheiden. Nordenfelt (2003, S. 1078) führt im Rahmen seiner Handlungstheorie an, dass sich eine Handlung aus „capacity aspects“, die hauptsächlich den Aktivitäten zuzuordnen sind und „performance aspects“, die die Partizipation widerspiegeln, zusammensetzt. Da eine Aktivität jedoch immer im Kontext der Umwelt zu betrachten sei, schlägt Nordenfelt (2003, S. 1078) eine Zusammenführung von Aktivitäten und Partizipation zur gemeinsamen Komponente „action“ vor.

Ewert, Allen, Wilson, Üstün und Stucki (2010) zeigen weiterhin, dass zwischen Aktivitäten und Partizipation ein hoher Zusammenhang besteht (Fragebogendaten von mehr als 3000 Patienten). Schuntermann

(2005, S. 6) präferiert hingegen eine Trennung beider Begriffe. Zum einen fehle der Partizipation die theoretische Basis, zum anderen entstehen damit unterschiedliche Fragestellungen.

In dieser Arbeit wird Mobilität dem Ansatz von Nordenfelt (2003) und Ewert et al. (2010) zugeordnet, da mobil zu sein nach WHO-Definition eine Interaktion mit der Umwelt impliziert.

Abbildung 3 veranschaulicht das ICF-Konzept sowie die Einordnung und Untergliederung der Mobilität auf erster Ebene.

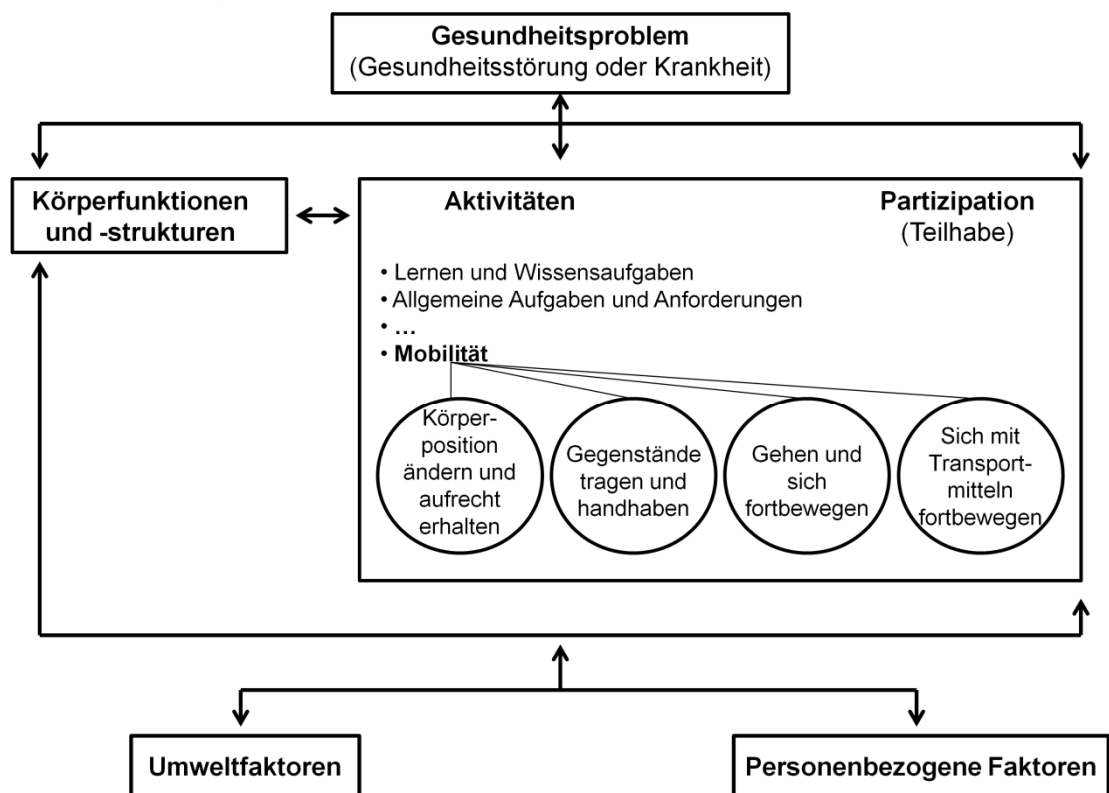


Abbildung 3: ICF-Modell mit Hervorhebung der Mobilitätskategorien (Kaczmarek et al., 2016a, S. 269)

Neben Bewegungshandlungen im Alltag weist die Mobilität auch einen Arbeitsbezug auf. So sollen sich beispielsweise rehabilitative Interventionen vermehrt an der Arbeitswelt der Patienten orientieren. Hierzu stellt die medizinisch-beruflich orientierte Rehabilitation (MBOR) ein mögliches Konzept dar (Müller-Fahrnow, Greitemann, Radoschewski, Gerwin, & Hansmeier, 2005). Dieses Konzept verfolgt eine inhaltlich individuelle Ausrichtung, die arbeitsplatzbezogene Elemente wie Stressbewältigung aber auch Mobilitätskategorien beinhaltet (Hillert, Koch, & Hedlund, 2007).

Mobilität im Sinne der ICF vereinigt weiterhin Bewegungshandlungen, die Voraussetzung zur Durchführung sportlicher Handlungen bzw. Teilnahme am Sport sind. Hierzu zählen beispielsweise die Mobilitätskategorien „Gegenstände mit den unteren Extremitäten bewegen“, „Werfen“, „Fangen“, „Klettern/Steigen“, „Rennen“, „Springen“, „Schwimmen“ (DIMDI, 2005, S. 106 f.).

Bezüglich der in Kapitel 2.1 aufgeführten Einflussfaktoren ist die Mobilität auch durch psychische Funktionen limitiert. Dies zeigt sich beispielsweise in der Bewegungsangst bei Rückenschmerzpatienten (Fear-Avoidance-Modell nach Pfingsten, 2004). So können die Mobilitätskategorien „Eine elementare Körperposition wechseln“ (d410) sowie „Gegenstände anheben und tragen“ (d430) durch psychosomatische Probleme beeinträchtigt sein. Patienten mit akuten Beschwerden versuchen häufig schmerzauslösende und möglicherweise verschlimmernde Bewegungen zu vermeiden. In diesem Zusammenhang wurde bereits in der Berliner Altersstudie nachgewiesen, dass die Mobilität ein wichtiger Prädiktor der Lebensqualität ist. Ein Mensch, der Strategien entwickelt, um gesundheitliche Probleme auf Körperfunktions- und -strukturebene zu kompensieren, ist in der Lage, Alltagshandlungen auszuführen und am gesellschaftlichen Leben teil zu haben (Smith, Fleeson, Geiselman, Settersten, & Kunzmann, 1996).

Solche Angstvermeidungsstrategien tragen demzufolge häufig zur kurzfristigen Linderung der Schmerzen bei. Demgegenüber steht die Entwicklung genereller Vermeidungsmuster bestimmter Aktivitäten, die sich langfristig negativ auf die Teilhabe auswirken können (Pfingsten, 2005). Diese Strategien können als personenbezogene Faktoren verstanden werden. Als Förderfaktoren sind sie in der Lage, Auswirkungen eines Gesundheitsproblems abzuschwächen. Negative Auswirkungen können aber auch das Problem verstärken (Barrieren) und folglich das soziale Leben beeinträchtigen (Kaczmarek et al., 2016a, S. 269).

In diesem Zusammenhang sind auch die ICF-Core-Sets zu nennen. Dabei handelt es sich um Auflistungen bestimmter ICF-Kategorien bzw. Funktionen, die für unterschiedliche Krankheitsbilder von Relevanz bzw. eingeschränkt sind (Stucki, Cieza, Ewert, Kostanjsek, Chatterji, & Üstün, 2002; Geyh et al., 2004).

Im Folgenden werden klassische Konzepte der Motorik beschrieben. Zudem erfolgt ein sprachlicher sowie inhaltlicher Vergleich zum Mobilitätskonzept.



## 2.3 Klassische Motorikkonzepte

Körperliche, willkürliche Bewegungen kommen durch im Körper stattfindende komplexe Prozesse zustande. Ausgehend von einer Bewegungsvorstellung und eines -ziels werden zunächst Areale im Gehirn aktiv (prämotorischer Kortex, Kleinhirn, Motorkortex, Hirnstamm), in denen die Bewegung räumlich und zeitlich geplant wird. Anschließend findet eine Informationsweiterleitung über Erregungsbahnen an das Rückenmark statt, das wiederum Kräfte- und Längenänderungen der entsprechenden Muskulatur aktiviert. Sensoren in den Muskeln (Muskelspindel) registrieren Längenveränderungen und melden Körperhaltung und/oder Bewegungen an das Zentrale Nervensystem zurück (De Marées, 2003, S. 68). Dieser sich ständig wiederholende Prozess wird nach sport-, insbesondere bewegungswissenschaftlichem Verständnis unter dem Begriff „Motorik“ zusammengefasst (Olivier, & Rockmann, 2003, S. 19; Roth, & Willimczik, 1999, S. 10 f.).

Zur Kategorisierung der Motorik hat sich das klassische Motorikmodell der Fähigkeiten und Fertigkeiten etabliert (Bös, & Mechling, 1983; Willimczik, & Singer, 2009, S. 18). In den Gesundheitswissenschaften kommen insbesondere ADL zur Anwendung (s. Kap. 2.1.1, S. 10). Eine neuere Herangehensweise ist das Modell der motorischen Kompetenzen (Herrmann, & Gerlach, 2014; Klafki, 2005; Pfeifer, Sudeck, Geidl, & Tallner, 2013). In diesem Zusammenhang erweitert die WHO mit dem Begriff Mobilität und dem damit verbundenen verstärkten Alltagsbezug das Motorikmodell (Kap. 2.2.2, S. 16).

Wie diese Modelle voneinander abzugrenzen sind bzw. welche Gemeinsamkeiten sich ergeben, wird im Folgenden erläutert.

### 2.3.1 Fähigkeiten und Fertigkeiten

Das Modell der motorischen Fähigkeiten nach Bös und Mechling (1983) ist in der sportwissenschaftlichen Theorie und Praxis seit Jahrzehnten etabliert. Fähigkeiten sind charakterisiert durch „individuelle Differenzen im Niveau der Steuerungs- und Funktionsprozesse, die bewegungsübergreifend von Bedeutung sind. Sie bilden die Voraussetzung für jeweils mehrere strukturell verschiedenartige Ausführungsformen und sind in ihrem Erklärungswert von unterschiedlicher Breite und Generalität“ (Roth, 1999, S. 233). Sie werden in konditionelle und koordinative Fähigkeiten sowie in „passive Systeme der Energieübertragung“ (Beweglichkeit) kategorisiert (Bös, & Mechling, 1983, S. 306; Abb. 4). Fähigkeiten stellen im Sinne

von Roth und Willimczik (1999) und Olivier und Rockmann (2003) motorische Prozesse dar und sind demnach nicht direkt beobachtbar.

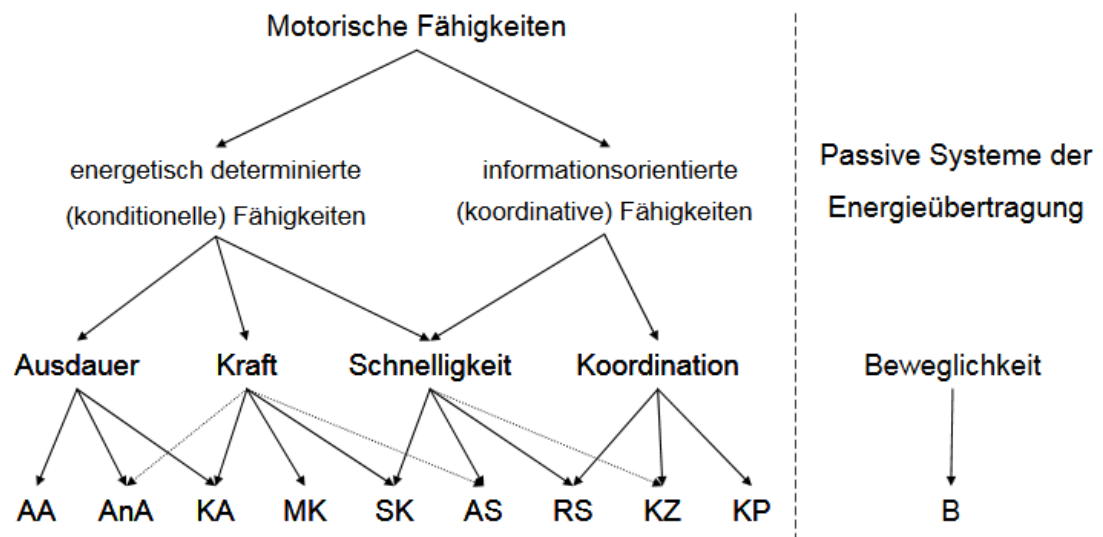


Abbildung 4: Motorische Fähigkeiten (Bös, 1987, S. 94)

Fertigkeiten werden nach Roth und Roth (2009, S. 227) wie folgt definiert: „Motorische Fertigkeiten kennzeichnen inter- und intraindividuelle Differenzen im Niveau der internen Steuerungs- und Funktionsprozesse, die der Realisierung einer spezifischen Bewegungstechnik zugrunde liegen. Sie sind sozusagen „eins zu eins“ mit einer äußerlich sichtbaren, strukturellen Ausführungsform verbunden“. Fertigkeiten sind demnach Bestandteil einer motorischen Handlung bzw. Bewegung und im Gegensatz zu Fähigkeiten direkt beobachtbar.

Bös und Mechling (1983, S. 67) unterschieden zwischen „Fertigkeiten des alltäglichen oder arbeits- bzw. sportspezifischen Verhaltens“. Burton und Miller (1998, S. 57) nahmen 15 Jahre später im Rahmen ihrer Taxonomie eine differenziertere Unterteilung in die Bereiche „locomotion“ (krabbeln, gehen, laufen, springen, usw.) und „object control“ (greifen, tragen, werfen, usw.) vor. Die Ausbildung erster Fertigkeiten wie rollen, krabbeln, sitzen, greifen, Gegenstände bewegen, sind in der frühen Kindheit verankert und werden als „early movement milestones“ beschrieben (Burton, & Miller, 1998, S. 179). Fertigkeiten, die einen aufrechten Stand erfordern (z. B. tragen, gehen, springen, laufen), werden im Laufe der Kindheit erworben und als „fundamental movement milestones“ charakterisiert (Burton, & Miller, 1998, S. 215). Aufbauend auf diesen elementaren Bewegungen können „sport skills“, Fertigkeiten, die zur Ausübung sportlicher Aktivitäten benötigt werden

(z. B. Korbwurf im Basketball), erworben werden (Burton, & Miller, 1998, S. 253).

In Anlehnung an diese Unterscheidung nehmen Roth und Roth (2009, S. 228) eine Unterteilung hinsichtlich der Komplexität einer Bewegung (elementar oder sportmotorisch) vor (Abb. 5). Weiterhin unterscheiden sie hinsichtlich der Variabilität. Dabei werden bestimmte Alltags- oder sportmotorische Fertigkeiten als „geschlossen“ charakterisiert, wenn sie unter gleichen Gegebenheiten und Bewegungsabläufen stattfinden (z. B. 100-m-Sprint). Fertigkeiten, die eine ständige Anpassung an die Situation erfordern, werden als „offen“ bezeichnet (z. B. Klettern; Roth, & Roth, 2009, S. 228).

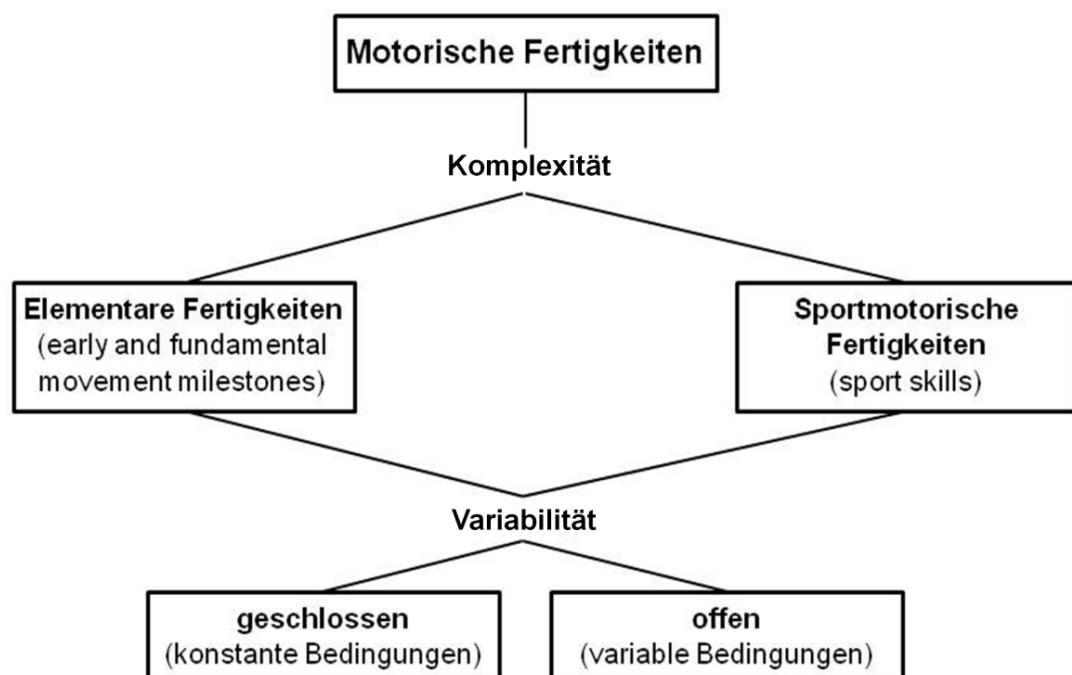


Abbildung 5: Motorische Fertigkeiten: Komplexitäts- und Variabilitätsgrad (modifiziert nach Roth, & Roth, 2009, S. 228)

Motorische Fähigkeiten werden erst durch die Ausübung von Fertigkeiten sichtbar und messbar. So liegt die terminologische Gleichsetzung der Begriffe Mobilität und Fertigkeiten sehr nahe (Wydra, & Kaczmarek, 2017, S. 116).

Motorische Fähigkeiten und Fertigkeiten werden hauptsächlich in der sportwissenschaftlichen Terminologie benutzt. Für andere Settings wie Bildungs- oder Gesundheitswissenschaften bildet das Konstrukt Kom-

petenz eine aktuelle Theoriebasis (Herrmann, & Gerlach, 2014; Pfeifer et al., 2013).

### 2.3.2 Kompetenzen

Der Begriff Kompetenz stammt in Deutschland aus den Bildungswissenschaften und basiert auf dem Klieme-Gutachten, das im Rahmen der Entwicklung nationaler Bildungsstandards entworfen wurde (Klieme et al., 2007). So finden sich auch im Kontext Schulsport neue Herangehensweisen und Modelle. Herrmann und Gerlach (2014) orientieren sich an motorischen Qualifikationen und Kompetenzen, um auch im Bereich Schulsport an die Diskussion anschlussfähig zu bleiben. Genauer unterscheiden sie zwischen „Motorischen Basisqualifikationen“ (MOBAQ) und Basiskompetenzen (MOBAK). MOBAQ sind hierbei im Sinne motorischer Fertigkeiten zu verstehen, während MOBAK übergeordnete Kategorien darstellen. Die Autoren orientieren sich am Ansatz von Burton und Miller (1998, S. 57; s. Kap. 2.3.1, S. 19) und unterteilen motorische Fertigkeiten in „Sich bewegen“ und „Etwas bewegen“. Das Zusammenspiel beider Fertigkeitsbereiche wird als „Motorische Kompetenz“ verstanden (Robinson et al., 2015, S. 1275). Der Ausprägungsgrad dieser motorischen Kompetenz beeinflusst dabei den Aktivitätsgrad im Alltag (Herrmann, & Gerlach, 2014).

Analog zu diesem Ansatz erfährt der Kompetenzbegriff aktuell auch innerhalb der Rehabilitation eine stärkere Bedeutung. Die Entwicklung von Kompetenzen dient hierbei dazu, die für eine Heilung einer Krankheit erforderlichen bewegungsbezogenen Aktivitäten ausführen und Alltagsanforderungen meistern zu können (Pfeifer, & Sudeck, 2016, S. 216).

Pfeifer et al. (2013, S. 12) verwenden hierbei den Begriff der „bewegungsbezogenen Gesundheitskompetenz“, die sich zusammensetzt aus

„kognitiven und motorischen Fähigkeiten und Fertigkeiten, die nötig sind, um gesundheitsförderliche körperliche Aktivität ausführen zu können, sowie aus den damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften bzw. Fähigkeiten zur erfolgreichen und kritisch-reflektierten Einbettung gesundheitsförderlicher körperlicher bzw. sportlicher Aktivität in variablen Situationen des Lebensalltags.“ (Pfeifer et al. 2013, S. 12 - 13)

Die Autoren beziehen sich auf die Aussagen von Baker (2006, S. 878), der das Verstehen gesundheitsrelevanter Informationen als Grundvoraussetzung zur Entwicklung eines gesundheitsorientierten Verhal-

tens auffasst. Abbildung 6 zeigt das von Pfeifer und Sudeck (2016) beschriebene Modell.

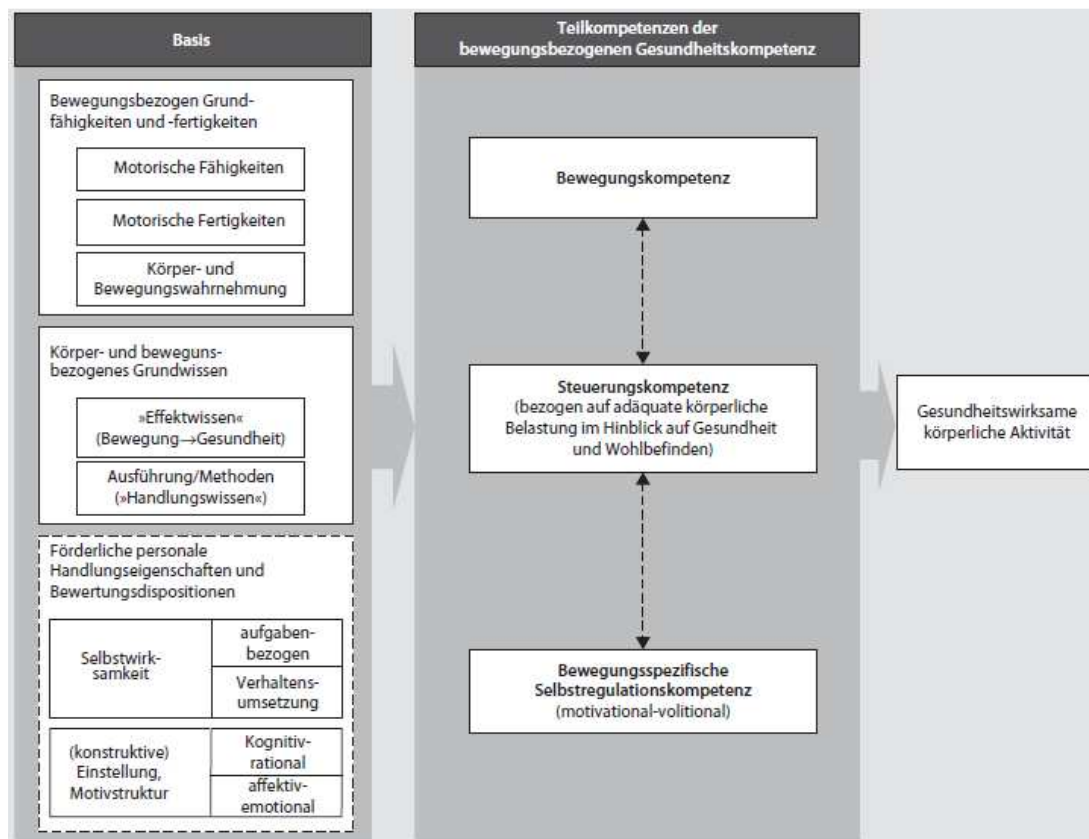


Abbildung 6: Bewegungsbezogene Gesundheitskompetenz (Pfeifer, & Sudeck, 2016, S. 222)

Die motorischen Fähigkeiten und Fertigkeiten, das Wissen über die gesundheitsförderliche Wirkung von Bewegung und mögliche Bewegungsausführungen, Selbstwirksamkeitserwartungen sowie die persönliche Motiv- und Motivationslage bilden die Grundlage dieser Kompetenz. Um eine Bewegungsform (z. B. schwimmen) auszuführen, sollte die Bewegung bereits in der Grobform beherrscht werden (Pfeifer, & Sudeck, 2016, S. 222).

Hummel und Borchert (2015, S. 5) weisen darauf hin, dass der Kompetenzbegriff keineswegs den Fähigkeitsbegriff verdrängen oder eine Alternative darstellen soll. Sie sollten aber auch nicht begrifflich synonym verwendet werden. Motorische Fähigkeiten und Fertigkeiten sind als wesentliche Elemente motorischer Kompetenzen zu betrachten. Die Autoren sprechen im Sinne der Fähigkeiten und Fertigkeiten auch von „motorischen Ressourcen“ (Hummel, & Borchert, 2015, S. 7), die getrennt von den Kompetenzen erfasst werden sollten.

Bezüglich der Operationalisierung motorischer Kompetenzen verweisen die Autoren auf ein noch bestehendes Forschungsdefizit. So können zwar im Setting Schule anhand der von Herrmann und Gerlach (2014) entwickelten MOBAKs motorische Qualifikationen (laufen, springen, werfen, usw.) erfasst werden. Diese lassen jedoch keine Rückschlüsse auf motorische Kompetenzen zu (Hummel, & Borchert, 2015, S. 6).

Unter motorischen Basiskompetenzen werden nicht direkt beobachtbare Voraussetzungen zur Absolvierung einer motorischen Leistung verstanden. Herrmann und Gerlach (2014) arbeiteten im Bereich Schulsport die motorischen Basisqualifikationen des „Sich bewege(n)s“ und „Etwas be- wege(n)s“ heraus. Inhaltlich vergleichbare Konstrukte seitens der Sport- wissenschaft und ICF können unter den Begriffen Fertigkeiten bzw. Mobilität verstanden werden (Wydra, & Kaczmarek, 2017, S. 118).

### **2.3.3 Aktivitäten des täglichen Lebens**

Es zeigt sich, dass die ICF mit der Aufführung der Mobilitätskategorien eine Vielzahl an Alltagstätigkeiten abdeckt (s. Abb. 3, S. 17). Hierbei stellt sich die Frage, inwieweit sich diese von den in den 1960er Jahren entwickelten ADL abgrenzen lassen.

Die Aktivitäten des täglichen Lebens bestehen u. a. aus den Kompo- nenten waschen, kleiden, essen und trinken sowie Transferbewegun- gen (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963, S. 915). Da die Mobilität im Sinne der ICF vergleichbare Tätigkeiten beinhaltet, können die ADL auch als Subkomponente der Mobilität angesehen werden. So werden unter Transferbewegungen beispielsweise die Fähigkeiten, un- abhängig zu Bett zu gehen und aus dem Bett zu steigen oder sich aus einer stehenden in eine liegende Position zu begeben, zusammenge- fasst. Die Mobilitätskategorien „Körperposition ändern und aufrecht er- halten“ und „Gehen und sich fortbewegen“ implizieren vergleichbare Tä- tigkeiten wie „Sich hinlegen“, „Sich beim Liegen verlagern“ und „Kurze Entfernungen gehen“ (Kaczmarek et al., 2016a, S. 271; Abb. 7).

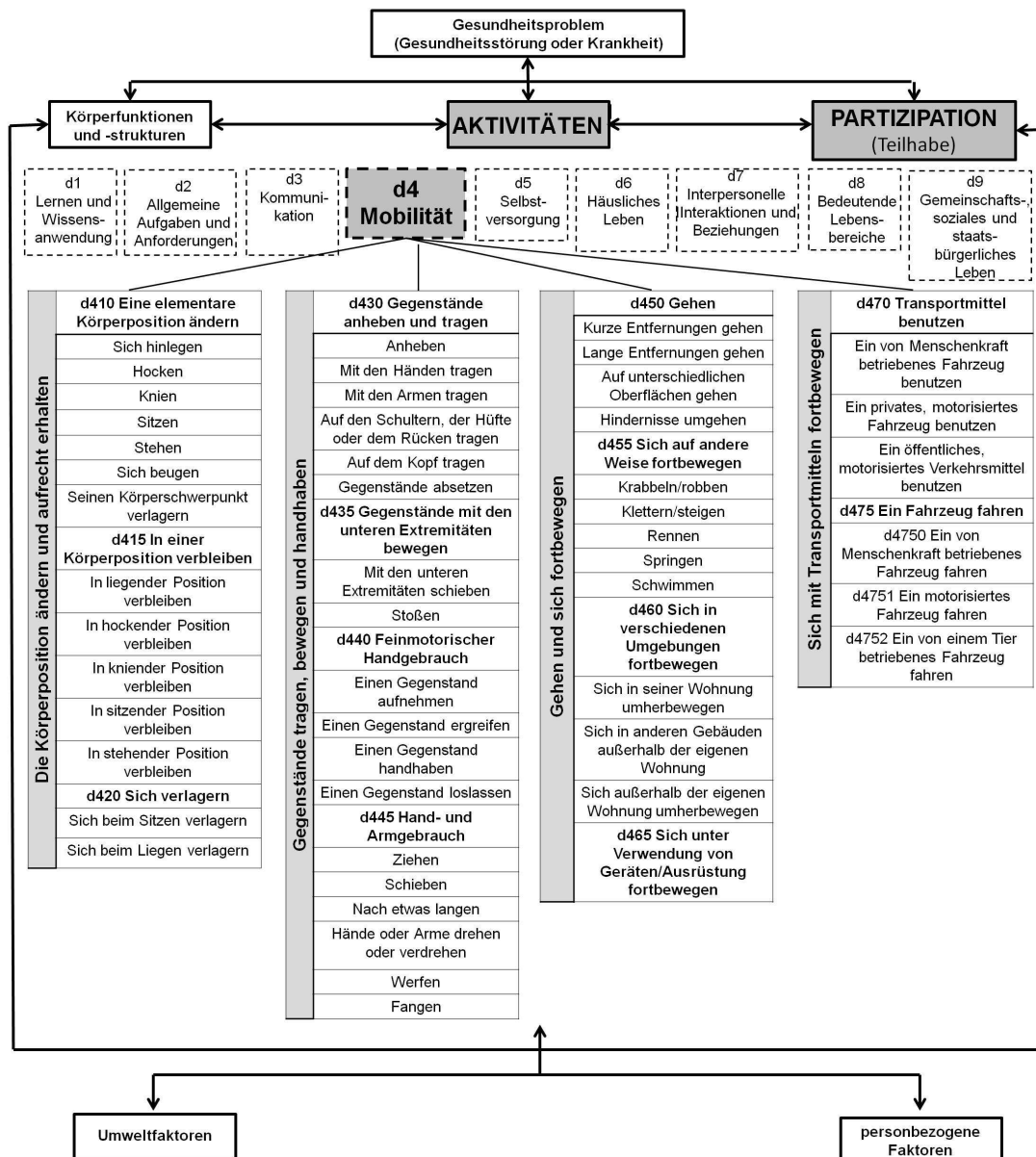


Abbildung 7: Unterkategorien der Mobilität innerhalb der ICF (modifiziert nach DIMDI, 2005, S. 23)

Ähnlich wie das ADL-Konzept umfasst die Mobilität mehrere Kategorien, die sich inhaltlich jedoch von den eigenständigen ICF-Domänen Selbstversorgung (d5) und Häusliches Leben (d6) abgrenzen (s. Abb. 7). Anhand der Abgrenzung wird deutlich, dass die Mobilität mit dem ADL-Konzept nicht gleichzusetzen ist. Mobilität meint mehr als die Komponente „Transferbewegungen“. Sie umfasst die in (Abb. 7) dargestellten vier Kategorien, die wiederum zusätzlich in Unterkategorien unterteilt sind.

## 2.4 Zusammenfassung

Mobil zu sein bedeutet, unabhängig mit der Umwelt zu interagieren. Diese Fähigkeit kann durch insbesondere alters- und krankheitsabhängige Faktoren negativ beeinflusst werden. Die WHO charakterisiert eine uneingeschränkte Mobilität hierbei als wichtiges Gesundheitskriterium und verankerte diese international in ihrer Klassifikation.

Die Mobilität bildet neben den klassischen Konzepten der motorischen Fähigkeiten und Fertigkeiten, Aktivitäten des täglichen Lebens und neueren Modellen der Kompetenzen eine weitere Betrachtungsweise der menschlichen Motorik. Sie beinhaltet Tätigkeiten, die zur Bewerkstellung alltäglicher Situationen notwendig sind. Dabei stellt sich die Frage, ob das Mobilitätskonzept im Sinne der ICF einen für die Sportwissenschaft kompatiblen, theoretischen Ansatz darstellt<sup>3</sup>.

Es zeigen sich begriffliche wie inhaltliche Gemeinsamkeiten zu den in der Literatur bekannten Aktivitäten des täglichen Lebens, aber auch zu Fertigkeiten und Kompetenzen (Tab. 3). Motorische Fähigkeiten können dabei als Voraussetzung zur Ausübung alltäglicher, sportlicher und beruflicher Tätigkeiten angesehen werden.

*Tabelle 3: Vergleichbare Motorikkonstrukte*

<b>Mobilität</b>	<b>Motorische Fertigkeiten</b>	<b>Motorische Basisqualifikationen</b>	<b>Aktivitäten des täglichen Lebens</b>
Körperposition ändern und aufrecht erhalten,  Gegenstände tragen und handhaben,  Gehen und sich fortbewegen,  Sich mit Transportmitteln fortbewegen	Gehen,  Laufen,  Greifen etc.	Prellen,  Dribbeln,  Fangen,  Werfen,  Balancieren,  Rollen,  Seitlaufen,  Springen	Transferhandlungen

<sup>3</sup> Die Diskussion um die Kompatibilität des Mobilitätsbegriffes wurde im Tagungsband *Bewegung, Raum und Gesundheit* veröffentlicht (Wydra, & Kaczmarek, 2017).



Schlussfolgernd können aus den angestellten Vergleichen und Abgrenzungen (s. Kap. 2.3) folgende Aspekte zur Mobilität festgehalten werden:

- Mobilität als Basis für jegliche gesundheitssportliche Aktivität bildet ein weiteres Konstrukt zur Beschreibung von motorischer Leistung und kann als Sammelbegriff für motorische Fertigkeiten verstanden werden.
- Aufgrund des durch die Weltgesundheitsorganisation hergestellten Kontextbezugs wird somit die motorische Leistungsfähigkeit in alltäglichen Situationen differenzierter erfassbar (Abb. 7, S. 24).

Die Orientierung am ICF-Modell der WHO eröffnet aus sportwissenschaftlicher Sicht eine Reihe von Anknüpfungspunkten. Neben möglichen Veränderungen von Therapieprogrammen stellen auch die Entwicklung und der Einsatz von Testverfahren wichtige Herausforderungen dar. Insbesondere im therapeutischen Kontext eignen sich hierbei ökonomische Methoden wie Screeningtests.

### 3 Screeningtests

Erfahrene Therapeuten erkennen häufig schon am Gangbild einer Person, ob motorische Einschränkungen vorliegen. Verfahren zur Beurteilung der Motorik können den „therapeutischen Blick“ absichern und gegebenenfalls spezifizieren. Insbesondere im klinisch-therapeutischen Kontext sind hierzu vor allem Methoden gefragt, die eine ökonomische Testung von Einzelpersonen, aber auch von Gruppen, ermöglichen. Motorische Screenings sollen dieses Kriterium erfüllen.

Im Folgenden wird zunächst der Assessmentprozess beschrieben. Im Weiteren folgt eine Gegenüberstellung von Tests und Screenings, die im klinisch-bewegungstherapeutischen Kontext Anwendung finden. Am Ende werden Besonderheiten von Evaluationsprozessen bei Screenings aufgeführt.

#### 3.1 Assessmentprozess

Zu Beginn der Rehabilitation liegen Informationen über eine Person bezüglich möglicher körperlicher, geistiger und motorischer Einschränkungen vor. Diese Informationen beruhen auf Ergebnissen verschiedener Tests aus z. B. Medizin, Sport-, Physio- oder Ergotherapie. Sie lassen Zielformulierungen zu, die als Grundlage zur Entscheidung über mögliche therapeutische Behandlungen genutzt werden können (Tack, 1976, S. 103).

Werden personenbezogene Informationen als nicht ausreichend bewertet, besteht die Möglichkeit weitere Tests durchzuführen („investigative Fragestellung“). Liegt hingegen eine terminale Entscheidung vor, können Behandlungen abschließend dazu genutzt werden den Therapieerfolg zu evaluieren und gegebenenfalls einen neuen diagnostischen Prozess einzuleiten (Tack, 1976, S. 104; Abb. 8).

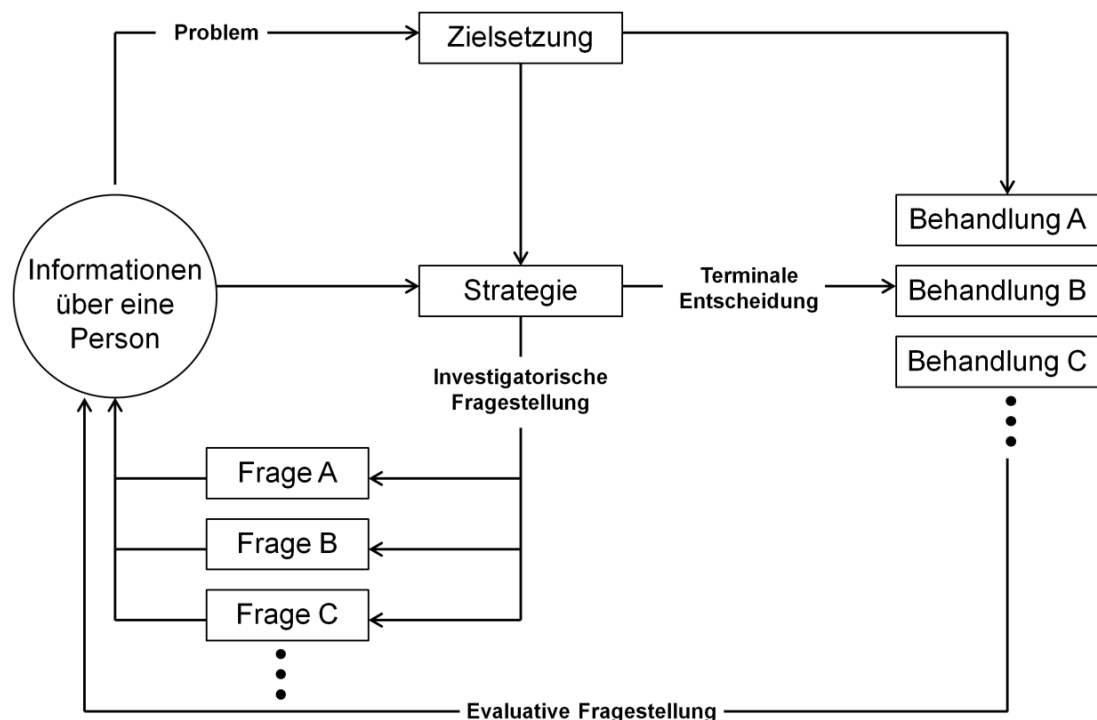


Abbildung 8: Diagnostischer Prozess (modifiziert nach Tack, 1976, S. 105)

In einem diagnostischen Prozess stellt die Gewinnung von Informationen zur Beurteilung der motorischen Leistungsfähigkeit (Bewegungstherapie) und bzw. oder zur Planung von Interventionen ein wesentliches Merkmal dar. Hierbei verweist Tack (1976, S. 104) darauf, dass Diagnostik nicht zum Selbstzweck betrieben werden, sondern mit möglichen Konsequenzen verbunden sein soll. Diagnostische Maßnahmen dienen hierbei als „Entscheidungshilfe“ (Tack, 1976, S. 104).

Klinische Einrichtungen sind nach § 135a Abs. 2 SGB V und § 20 Abs. 2 SGB V zur Anwendung qualitätssichernder Maßnahmen verpflichtet (DRV, 2015a, S. 655, 1520). Um die Akzeptanz durch Kostenträger zu gewährleisten, sollen diese vermehrt auf der Basis der ICF erfolgen (Baldus, Schüle, & Huber, 2012, S. 326). Entsprechend ausgerichtete Assessmentverfahren eignen sich hier, um beispielsweise die Qualität und Zielerreichung bewegungstherapeutischer Interventionen (Förderung körperlicher Gesundheit) zu evaluieren. Wydra (2006a, S. 10) charakterisierte den Begriff Assessment wie folgt:

- Im Gesundheitsbereich werden Assessments für den gesamten Komplex der interdisziplinären Anamnese-, Befunderhebung, Krankenbeobachtung, das Sammeln und Bewer-

ten von Informationen aus verschiedensten Quellen und das Formulieren diagnostischer Arbeitshypothesen verwendet.

- Assessment bedeutet in dem Fall ein multidimensionaler und interdisziplinärer Prozess, mit dem Ziel einen Behandlungs- und Betreuungsplan zu entwickeln (Wydra, 2006a, S. 10).

### 3.2 Sequentielle Diagnosestrategie

Eine optimale Diagnostik erfordert einen maximalen Informationsgewinn bei gleichzeitig möglichst ökonomischem Testaufwand. Die Integration von Screenings in sequentielle Diagnosestrategien führt meist zu einem Ökonomisierungsprozess (Bös, Wydra, & Karisch, 1992, S. 110; Wydra, 2006a, S. 16).

Nach medizinischer Diagnose soll die Frage beantwortet werden, ob Einschränkungen hinsichtlich der Motorik vorliegen und welche Therapieform die jeweils geeignete darstellt (Abb. 9).

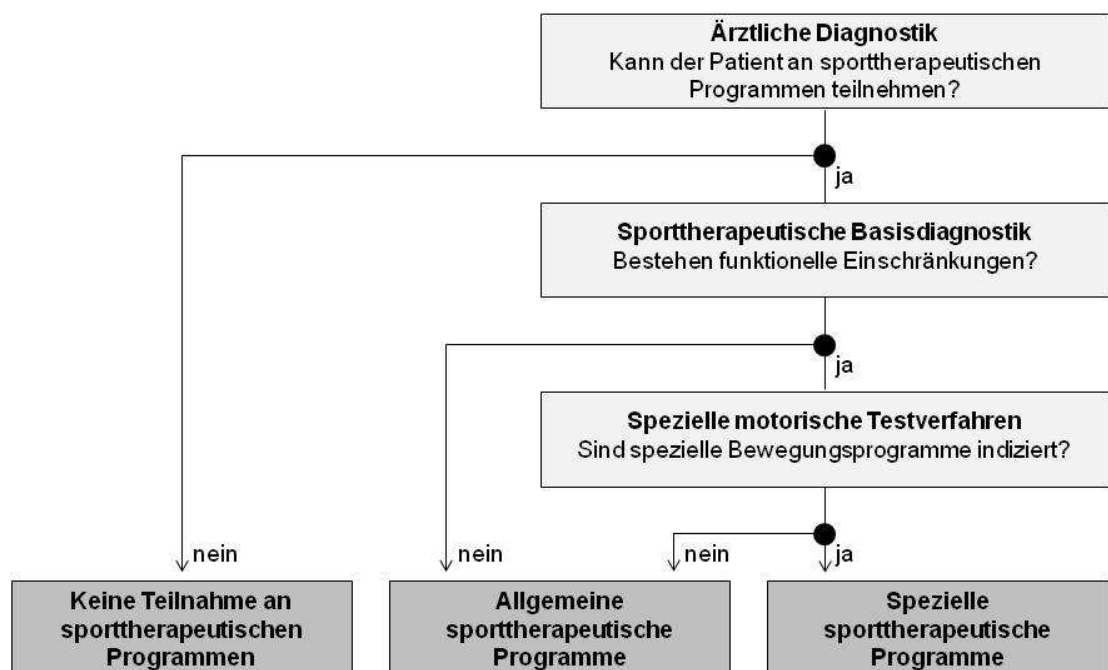


Abbildung 9: Sequentielle Diagnosestrategie (modifiziert nach Wydra, 2006a, S. 17)

Screeningverfahren können hierbei auf der zweiten Stufe (Sporttherapeutische Basisdiagnostik) eingesetzt werden, um einen ersten Eindruck über die motorische Leistungsfähigkeit bzw. Mobilität zu erhalten. Bestehen motorische Einschränkungen, können diese in weiteren Tests spezifiziert werden. Durch den Einsatz von Screenings kann der Diagnoseprozess möglicherweise optimiert bzw. verkürzt werden, da Personen mit unauffälligem Befund nicht weiter diagnostiziert werden müssen. Mariush (2004, S. 63) empfiehlt bei Personen mit motorischen Einschränkungen die Anwendung von stärker differenzierenden motorischen Tests und gegebenenfalls eine weitergehende ärztliche Untersuchung.

### 3.3 Allgemeine Test- vs. Screeningverfahren

Zum Diagnostizieren motorischer Auffälligkeiten wird im allgemeinen Sprachgebrauch häufig der Begriff „Test“ verwendet. Lienert und Raatz (1998, S. 7) beschreiben einen Test als „wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung“.

Tests finden nach Ballreich (1970, S. 19 ff.) in den folgenden Bereichen Anwendung:

*Leistungsdiagnostik:* Zur Überprüfung der Leistung kommen Testverfahren zum Einsatz, die eine Beurteilung motorischer Fähigkeiten und Fertigkeiten, aber auch taktischer Fähigkeiten einer Person zulassen.

*Entwicklungsdiagnostik:* Durch die zweimalige Anwendung eines Tests zu zwei verschiedenen Zeitpunkten kann eine Veränderung der Leistung überprüft werden.

*Eignung und Talent:* Tests, die Prognosen hinsichtlich anthropometrischer und leistungsbezogener Bedingungen einer Person zulassen, ermöglichen „Talente“ für bestimmte Sportarten herauszufiltern.

*Dimensionsanalyse:* Diagnostische Methoden, mittels derer Teilkomponenten einer Bewegung erfasst werden können, tragen zur Optimierung der Gesamtbewegung bei.

*Experimenteller Bereich:* Durch Änderung der Trainingsmethodik lassen sich Auswirkungen auf die Motorik messen.

Für die Sport- und Bewegungstherapie sieht Ludwig (2012, S. 175) vor allem die Leistungs-, Entwicklungs- und Eignungsdiagnostik als zielführend. Mittels motorischer Tests erhält der Therapeut einen Eindruck über die motorische Leistungsfähigkeit einer Person (Leistungsdiagnos-

tik). Hier spielt insbesondere die Normierbarkeit des Tests eine Rolle. Entwicklungsdiagnostische Verfahren kommen zum Einsatz, um Veränderungen der motorischen Leistungsfähigkeit durch Reifungsprozesse – wie sie bei kleinen Kindern zu beobachten sind – oder therapeutische Interventionen – z. B. in der Sporttherapie – zu erfassen. Hier spielt die Responsivität des Testverfahrens eine wesentliche Rolle.

Im Gegensatz zu Tests, die meist eine Überprüfung mehrerer Persönlichkeitsmerkmale erlauben, zielen Screenings auf die *Dichotomisierung* eines Merkmals ab. Dabei soll ein Merkmal möglichst mit zwei extremen Gegenpolen charakterisiert werden (z. B. auffällig vs. nicht auffällig; krank vs. gesund; Wydra, & Kaczmarek, 2016, S. 84). Mariush (2004, S. 63) bezeichnet ein Screening daher plakativ als Sieb, in dem Personen mit Krankheiten (positives Testergebnis) von Personen ohne Krankheiten (negatives Testergebnis) getrennt werden:

- Screeningtests erlauben eine grobe Identifizierung motorischer Einschränkungen bei größeren Stichproben.
- Weiterführende Testverfahren ermöglichen im Anschluss eine Differenzierung individueller motorischer Auffälligkeiten bei in der Regel reduzierten Stichprobengrößen.

Zur Charakterisierung von Screenings wird bis heute die von der United States Commission of Chronic Illness aufgestellte Definition verwendet (Wilson, Jungner, & WHO, 1968, S. 11):

„The presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not“. (Commission of Chronic Illness, 1957, zitiert nach Wilson et al., 1968, S. 11)

Die bestehende Definition wurde in der Vergangenheit erweitert. So betonen Croyle (1995, S. 3) die Bedeutung der Testökonomie und sehen den Vorteil in ihrer Einfachheit. Neben der ökonomischen Anwendbarkeit sollen Screenings auch tatsächlich eine Verbesserung der Entdeckungsrate ermöglichen (Croyle, 1995, S. 3, Mariush, 2004, S. 63). Weiterhin können sie nicht nur zur Identifikation bereits bestehender Krankheiten, sondern vor allem zur Erfassung krankheitsrelevanter Risikofaktoren angewendet werden, die wiederum den Einsatz präventiver Maßnahmen legitimieren.

### 3.3.1 Evaluationskriterien

Screeningtests müssen wie allgemeine Tests unter standardisierten Bedingungen ablaufen und somit den Haupttestgütekriterien genügen. Darüber hinaus entscheidet insbesondere das Nebengütekriterium „Ökonomie“, ob ein Verfahren in der Praxis Anwendung findet (Bös, 2017, S. 826). So sollen Screenings schnell und einfach durchführbar sein und sich am praktischen Nutzen orientieren (Wilson et al., 1968, S. 23).

Im Sinne prognostischer Validität soll ein Screening in der Lage sein, Auffälligkeiten (z. B. motorischer Art) zielsicher zu erkennen oder auszuschließen (Knottnerus, & Muris, 2002, S. 3). Deshalb ist neben der Beurteilung der üblichen Gütekriterien die diagnostische Qualität von besonderer Bedeutung (Fryback, & Thornbury, 1991, S. 88; Sackett, & Haynes, 2002, S. 539). Beide Gütemaße sind in Abbildung 10 dargestellt.

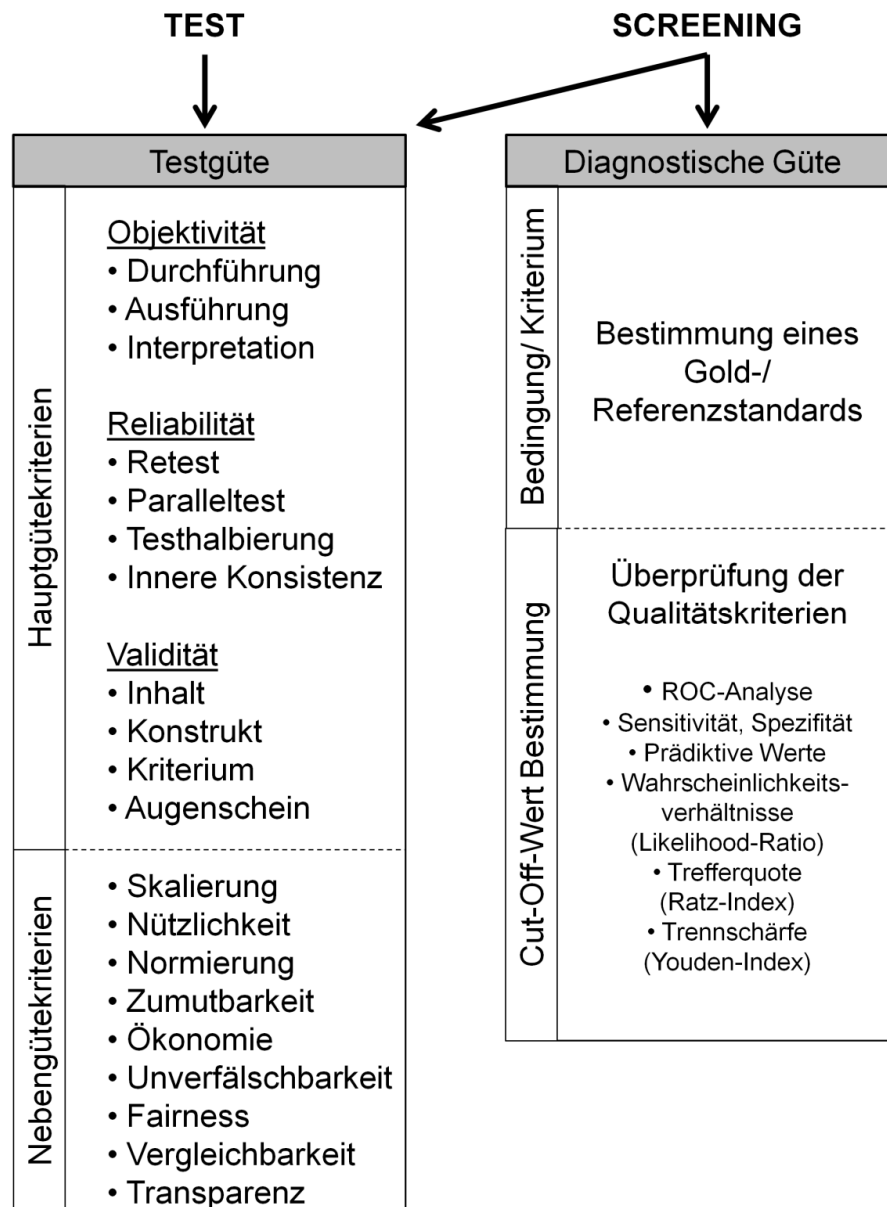


Abbildung 10: Evaluationskriterien für allgemeine Tests und Screeningtests

Die diagnostische Güte spielt vor allem bei der Evaluation medizinischer Verfahren eine wichtige Rolle (Leeftang, Rutjes, Reitsma, Hooft, & Bossuyt, 2013, S. 537). So wurde beispielsweise die Qualität der Mammographie permanent weiterentwickelt, da ein zielsicheres Erkennen eines Mammakarzinoms lebensentscheidend sein kann (Berg et al., 2004; Evans et al., 2001; Pisano et al., 2008).



### 3.3.2 Testgüte und diagnostische Güte motorischer Tests und Screenings

In der Bewegungstherapie kommen unterschiedliche Screenings zur Anwendung. Tabelle 4 gibt einen Überblick an motorischen Verfahren, die insbesondere in der klinischen Praxis und in Studien häufig verwendet werden (Bennell, Dobson, & Hinman, 2011; Harada, Chiu, Damron-Rodriguez, Fowler, Siu, & Reuben, 1995; Perell, Nelson, Goldman, Luther, Prieto-Lewis, & Rubenstein, 2001; Scharhag-Rosenberger et al., 2014).

- *Timed Up and Go Test* (TUG; Podsiadlo, & Richardson, 1991)  
Der TUG wurde konzipiert, um die Mobilität von Personen in der Geriatrie beurteilen zu können. Weiterhin wird er zur Einschätzung des Sturzrisikos verwendet.
- *Berg-Balance-Scale* (Berg, Wood-Dauphinee, Williams, & Gayton, 1989)  
Anhand einer Skala, bestehend aus 14 Items, soll das Sturzrisiko bei älteren Menschen beurteilt werden.
- *Chair Stand Test* (McMurdo, & Rennie, 1993)  
Dieser Test dient zur Beurteilung der Beinkraft.
- *Tinetti Test* (Tinetti, Williams, & Mayewski, 1986)  
Der Tinetti Test ist ein Testverfahren zur Beurteilung des Sturzrisikos.
- *6 Minute Walking Test* (Butland, Pang, Gross, Woodcock, & Geddes, 1982)  
Dieser Test soll anhand der Gehfähigkeit Aussagen über die funktionale Leistungsfähigkeit ermöglichen. Gerade bei der Evaluation therapeutischer Programme von Lungen- und Herzpatienten kommt dieser Test zum Einsatz.

Die diagnostische Güte der aufgeführten Tests wurde hauptsächlich hinsichtlich ihrer Vorhersagbarkeit des Sturzrisikos überprüft (Tab. 4).

*Tabelle 4: Auswahl motorischer Screenings: Studien zur Testgüte und diagnostischen Güte*

Screening	Testgüte	Diagnostische Güte
<b>Timed Up and Go Test</b> (Podsiadlo, & Richardson, 1991)	<ul style="list-style-type: none"> <li>- Interrater-Reliabilität (Podsiadlo, &amp; Richardson, 1991; Shumway-Cook et al., 2000; Morris et al., 2001; Schoppen et al., 1999; Nordin et al., 2006)</li> <li>- Test-Retest-Reliabilität (Rockwood et al., 2000; Huang et al., 2011; Flansbjer et al., 2005; Ries et al., 2009; Ng, &amp; Hui-Chan, 2005)</li> <li>- Validität (Podsiadlo, &amp; Richardson, 1991; Schoppen et al., 1999; Bennie et al., 2003; Lin et al., 2004)</li> </ul>	<ul style="list-style-type: none"> <li>- Sturzrisiko älterer Personen (Shumway-Cook et al., 2000; Rose et al., 2002; Chiu et al., 2003; Lin et al., 2004; Whitney et al., 2004; Arnold, &amp; Faulkner, 2007; Kristensen et al., 2007; Nordin et al., 2008; Haines et al., 2008; Shimada et al., 2009; Alexandre et al., 2012; Kojima et al., 2015)</li> <li>- Mobilität älterer Personen (Bischoff et al., 2003; Savva et al., 2012)</li> </ul>
<b>Berg-Balance-Scale</b> (Berg, Wood-Dauphinee, & Williams, 1995)	<ul style="list-style-type: none"> <li>- Interrater-Reliabilität (Berg et al., 1995; Mao et al., 2002)</li> <li>- Intrarater-Reliabilität (Berg et al., 1995)</li> <li>- Interne Konsistenz (Berg et al., 1995; Mao et al., 2002; Chou et al., 2006)</li> <li>- Kriterienvvalidität (Mao et al., 2002; Chou et al., 2006)</li> <li>- Konstruktvalidität (Mao et al., 2002)</li> </ul>	<ul style="list-style-type: none"> <li>- Sturzrisiko (Scale, &amp; Talking, 2006; Alzayer et al., 2009)</li> </ul>
<b>Chair Stand Test</b> (McMurdo, & Rennie, 1993)	<ul style="list-style-type: none"> <li>- Interrater-Reliabilität (Almeida et al., 2010; Wright et al., 2011)</li> <li>- Kriteriumsvalidität (Jette et al., 1999; Almeida et al., 2010)</li> </ul>	<ul style="list-style-type: none"> <li>- Sturzrisiko (Lipsitz et al., 1991)</li> </ul>
<b>Tinetti Performance-Oriented Mobility Assessment</b> (Tinetti et al., 1986)	<ul style="list-style-type: none"> <li>- Intrarater-Reliabilität (Kegelmeyer et al., 2007)</li> <li>- Interrater-Reliabilität (Faber et al., 2006; Kegelmeyer et al., 2007)</li> <li>- Test-Retest Reliabilität (Faber et al., 2006; Canbek et al., 2013)</li> <li>- Kriteriumsvalidität (Faber et al., 2006; Canbek et al., 2013)</li> </ul>	<ul style="list-style-type: none"> <li>- Sturzrisiko (Faber et al., 2006; Kegelmeyer et al., 2007; Panzer et al., 2011)</li> </ul>
<b>Six-Minute Walk Test</b> (Butland et al., 1982)	<ul style="list-style-type: none"> <li>- Test-Retest Reliabilität (Demers et al., 2001; Kennedy et al., 2005; Collins et al., 2008)</li> <li>- Interne Konsistenz (Collins et al., 2008)</li> <li>- Kriteriumsvalidität (Demers et al., 2001; Collins et al., 2008)</li> </ul>	<ul style="list-style-type: none"> <li>- Prognose von Herzfehlern (Rubim et al., 2006)</li> </ul>

### 3.3.3 Diagnostische Güte

Die gewünschte Dichotomisierung (s. Kap. 3.3, S. 31) macht einen Trennpunkt (Cut-Off-Wert) notwendig, um auffällige (positive) oder unauffällige (negative) Ergebnisse eindeutig bewerten zu können. Hierzu ist zunächst die Wahl eines optimalen Gold- bzw. Referenzstandards notwendig.

#### Gold- bzw. Referenzstandard

Das Screening wird mit einem sogenannten „Goldstandard“ verglichen. Dabei handelt es sich um eine Methode bzw. einen Test, die bzw. der Personen hinsichtlich ihrer Auffälligkeit möglichst sicher identifizieren kann (Reitsma, Rutjes, Khan, Coomarasamy, & Bossuyt, 2009, S. 797). Da in der Praxis ein solcher Standard oft nicht gegeben ist, wird auf die Methode zurückgegriffen, die den Zustand am zielsichersten beurteilen kann (Referenzstandard; Bossuyt et al., 2003, S. 827; Knottnerus, & Muris, 2002, S. 51). Manche Autoren sprechen hierbei auch von einem „Indextest“ (Bossuyt et al., 2015, S. 828).

Im Vergleich mit dem Referenzstandard sind folgende Ergebnisse möglich (Akobeng, 2007, S. 339):

- **Richtig-positiv:** Bestätigt die am Referenzstandard orientierte Auffälligkeit.
- **Falsch-positiv:** Ergibt ein positives Ergebnis, obwohl der Referenzstandard keine Auffälligkeit anzeigt.
- **Falsch-negativ:** Ergibt ein negatives Ergebnis, obwohl der Referenzstandard eine Auffälligkeit erkennt.
- **Richtig-negativ:** Bestätigt das durch den Referenzstandard gefundene unauffällige Ergebnis.

Durch diese vier Optionen wird eine eindeutige Beurteilung des Ergebnisses gewährleistet. Es können weitere diagnostische Maßnahmen getroffen bzw. die Wahl geeigneter Therapieansätze erleichtert werden (Knottnerus, Buntinx, & Van Weel, 2009, S. 3).

Richtlinien zur Durchführung und Berichterstattung diagnostischer Studien bieten die Standards for Reporting Diagnostic Accuracy (STARD; Bossuyt et al., 2015). Zur Berichterstattung sind neben der Wahl eines optimalen Referenzstandards auch entsprechende Qualitätskriterien<sup>4</sup> von Bedeutung.

---

<sup>4</sup> Die Auswahl begrenzt sich auf die in dieser Arbeit verwendeten Kriterien.

## Qualitätskriterien

### ROC-Analyse

Zunächst wird überprüft, inwiefern das Screening in der Lage ist, die Dichotomisierung richtig darzustellen. Hierzu werden Sensitivität und Spezifität bestimmt.

- *Sensitivität*: Die Richtig-Positiv-Rate gibt Aufschluss darüber, wie viele Personen der tatsächlich körperlich Beeinträchtigten durch das Screening als solche erkannt werden (Zhou et al., 2002, S. 16).
- *Spezifität*: Die Richtig-Negativ-Rate gibt an, wie viele Personen der tatsächlich körperlich Unbeeinträchtigten durch das Screening als solche klassifiziert werden (Zhou et al., 2002, S. 16).

Weiterhin können prädiktive Werte ermittelt werden.

- *Positiver Prädiktiver Wert*: Dieser Wert zeigt, mit welcher Wahrscheinlichkeit Personen mit einem positiven Testergebnis tatsächlich an körperlichen Krankheiten, Einschränkungen oder Verletzungen leiden (Zhou et al., 2002, S. 44 f.).
- *Negativer Prädiktiver Wert*: Dieser Wert bedeutet die Wahrscheinlichkeit, dass Personen mit negativem Testergebnis tatsächlich nicht an körperlichen Beeinträchtigungen leiden (Zhou et al., 2002, S. 45).

Abb. 11 veranschaulicht das Zustandekommen von Sensitivität, Spezifität und der prädiktiven Werte anhand des Vergleichs der Ergebnisse von Referenzstandard und Screening (Vierfeldertafel). Die Sensitivität wird dabei bestimmt durch das Verhältnis zwischen richtig-positiven und falsch-negativen Befunden. Hinsichtlich der Spezifität werden falsch-positive und richtig-negative Ergebnisse betrachtet. Positive prädiktive Werte ergeben sich aus dem Verhältnis richtig-positiver und falsch-positiver Befunde. Falsch-negative und richtig-negative Ergebnisse werden zur Bestimmung negativer, prädiktiver Werte verwendet.

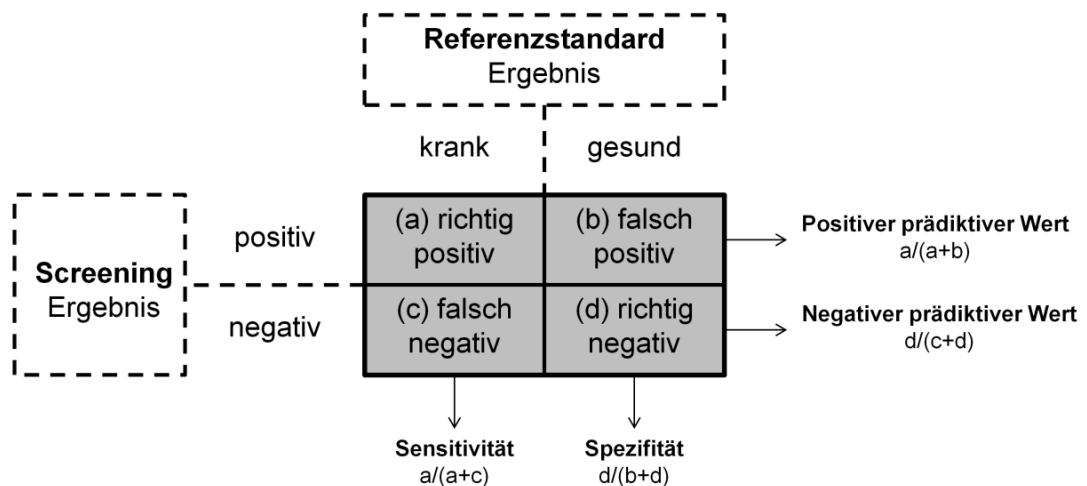


Abbildung 11: Vierfeldertafel

Nach Bestimmung von Sensitivität und Spezifität kann ein möglicher Trennpunkt für auffällige und nicht auffällige Testergebnisse bestimmt werden. Um jedem Testwert eines Screenings einen Sensitivitäts- und Spezifitätswert zuzuordnen, wird hierzu ein Kurvendiagramm verwendet (Receiver Operating Characteristic Curve – ROC; Hanley, & McNeil, 1982; Zou, O'Malley, & Mauri, 2007; Abb. 12).

Die Winkelhalbierende (C: gestrichelte Linie) zeigt die Trefferquote, die durch eine zufällige Urteilsfällung resultieren würde. Befindet sich die Kurve (B) oberhalb dieser Linie, besitzt das Screening eine höhere Trefferquote im Vergleich zum Zufall. Dies wird auch durch die Fläche unterhalb der Kurve (Area Under The Curve – AUC) dargestellt.

Ein AUC-Wert von 0,5 bedeutet, dass die Trefferquote des Screenings mit einer Trefferquote durch den Zufall gleichzusetzen ist (Abb. 12, Winkelhalbierende). Bei einer AUC von 1 besteht eine 100-prozentige Sensitivität und Spezifität. Die Wahl des Cut-Off-Wertes orientiert sich hierbei an dem Punkt, der am weitesten von der Winkelhalbierenden entfernt bzw. am nächsten zum Optimalpunkt (A: AUC = 1) liegt (Zou et al., 2007, S. 655).

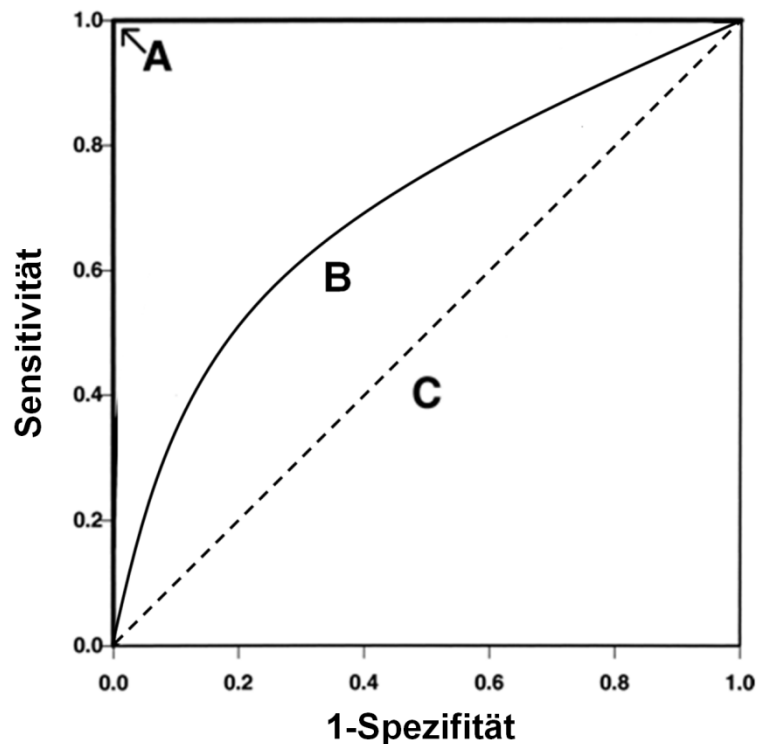


Abbildung 12: Hypothetische Verhältnisannahme von Sensitivität und Spezifität und daraus resultierende ROC. A:  $AUC = 1$  (optimale Trefferquote: 100 % Sensitivität und Spezifität); B:  $AUC \approx 0,7$  (höhere Trefferquote des Screenings als Zufallstrefferquote); C:  $AUC = 0,5$  (keine höhere Trefferquote im Vergleich zur Zufallstrefferquote; modifiziert nach Zou et al., 2007, S. 655)

Weiterhin ermöglicht die Verwendung der Trennschärfe (Youden-Index) eine ökonomische Bestimmung eines optimalen Cut-Off-Wertes.

#### *Trennschärfe (Youden-Index)*

Der Youden-Index (YI;  $\text{Sensitivität} + \text{Spezifität} - 1$ ) gibt Aufschluss über die Effektivität eines Screenings und kann als Trennschärfekoeffizient angesehen werden (Fluss, Faraggi, & Reiser, 2005; Youden, 1950). Ein optimaler Cut-Off-Wert liegt in dem Punkt, in dem sich Sensitivität und Spezifität in bestmöglichem Gleichgewicht befinden (Summe aus höchstem Sensitivitäts- und Spezifitätswert).

Der Index liegt zwischen 0 und 1. Je näher der Index an 1 liegt, desto besser ist das Screening in der Lage zwischen „körperlich unbeeinträchtigten“ und „beeinträchtigten“ Personen zu unterscheiden (Fluss et al., 2005, S. 458).

### Wahrscheinlichkeitsverhältnisse (Likelihood-Ratio)

Ein weiteres Maß zur Beurteilung der Qualität eines Screenings bieten die Wahrscheinlichkeitsverhältnisse. Der Likelihood-Ratio bezeichnet das Wahrscheinlichkeitsverhältnis für ein positives oder negatives Testergebnis (Zhou et al., 2002, S. 35 ff.; Tab. 5).

*Tabelle 5: Diagnostische Evidenz in Abhängigkeit der Wahrscheinlichkeitsverhältnisse (Jaeschke, Guyatt, & Sacket, 1994, zitiert nach Denegar, & Fraser, 2006, S. 204)*

<b>Positive Likelihood-Ratio (LR+)</b> Sensitivität / 1-Spezifität	<b>Negative Likelihood-Ratio (LR-)</b> 1-Sensitivität / Spezifität	<b>Diagnostische Evidenz</b>
> 10	< 0,1	Hoch
5 - 10	0,1 - 0,2	Moderat
2 - 5	0,2 - 0,5	Gering
1 - 2	0,5 - 1	Sehr gering

Der positive Likelihood-Ratio (LR+) zeigt, wie wahrscheinlicher ein positives Testergebnis bei körperlich Beeinträchtigten ist im Vergleich zu Unbeeinträchtigten. Der negative Likelihood-Ratio (LR-) gibt an, wie viel wahrscheinlicher es ist, dass körperlich Beeinträchtigte ein negatives Testergebnis erzielen im Vergleich zu Unbeeinträchtigten.

### Trefferquote (Ratz-Index)

Neben der AUC ist der Ratz-Index (RI; Trefferquote minus Zufallstrefferquote / Max. Trefferquote minus Zufallstrefferquote) ein weiteres Maß zum Vergleich der Trefferquote des Screenings mit der Zufallstrefferquote. Werte ab 0,6 bedeuten eine große Verbesserung (Marx, & Lenhard, 2010, S. 72 f.).

### 3.3.4 Grenzwertproblematik

Die Setzung eines Norm- bzw. Grenzwertes erfolgt entweder auf der Grundlage von Erfahrungen, die im Laufe von Testanwendungen gesammelt wurden oder anhand aufgestellter Standards bzw. Berechnungen. Je nach Setting (z. B. im Gesundheits- und Bildungssektor) existieren verschiedene Normen, Standards und Verfahren, die zur Legitimierung eines Grenzwertes herangezogen werden (Tab. 6).

*Tabelle 6: Grenzwertbestimmungen (modifiziert nach Israel, 1983; KMK, 2005; Wydra, 2006b)*

Gesundheitsbereich			Bildungsbereich
Normen	Diagnostische Güte	ICF	Standards
Majoritätsnormen	Cut-Off-Werte (Sensitivität, Spezifität, usw.)	Beurteilungs- merkmale – Ausmaß des Problems	Mindest-/ Minimalstandards
Minimalnormen			Regelstandards
Idealnormen			Exzellenz-/ Maximalstandards
Spezialnormen			

## Normen

Israel (1983) unterschied zwischen Majoritäts-, Minimal-, Ideal- und Spezialnormen:

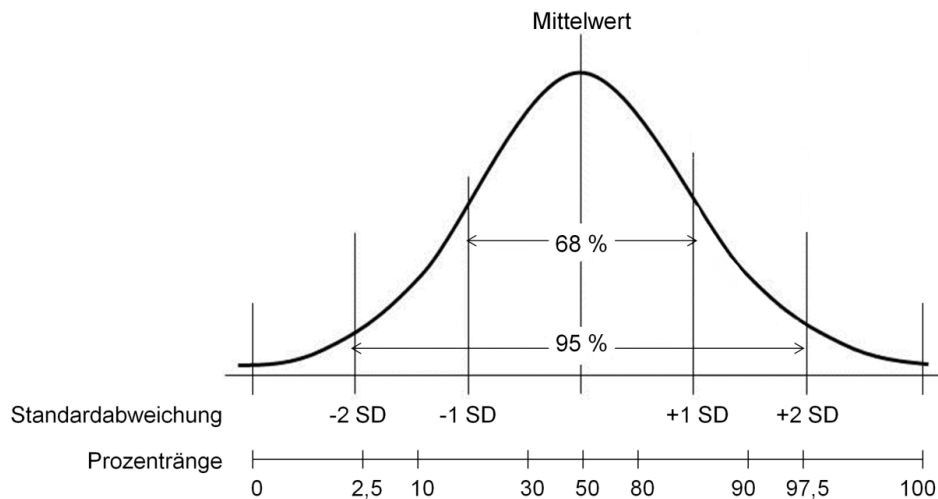
### *Majoritätsnormen*

Häufig werden Grenzwerte anhand des Mittelwerts der Testergebnisse einer Stichprobe bestimmt. Bei großen Stichproben, deren Ergebnisse normalverteilt sind, geben Prozentränge Aufschluss über auffällige bzw. unauffällige Resultate. Hierbei werden Ergebnisse, die weniger als eine Standardabweichung oberhalb und unterhalb des Mittelwerts liegen (68 %) häufig als unauffällig klassifiziert (Israel, 1983; Wydra, 2006b). Weiterhin kann auch das 95%-Konfidenzintervall (zwei Standardabweichungen) herangezogen werden (Abb. 13). Je nach angelegtem Maß verschieben sich dann mögliche Grenzwerte.

Bei nicht normalverteilten Ergebnissen können Grenzwerte anhand eines Boxplots bestimmt werden. Hierbei werden Minimum- und Maximumwerte, Median, unteres und oberes Quartil und Interquartilsabstand bestimmt.



a)



b)

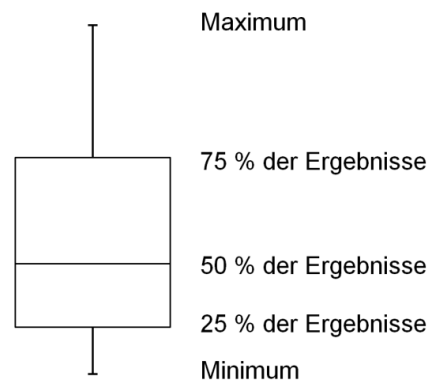


Abbildung 13: Grenzwertbestimmung bei a) normal- und b) nicht normalverteilten Ergebnissen

Majoritätsnormen, die anhand der erzielten Mittelwerte einer Referenzstichprobe ermittelt wurden, sind dementsprechend lediglich für diese Personengruppe repräsentativ. Die Anwendung eines motorischen Screenings mit Patienten unterschiedlicher Krankheitsbilder erschwert hierbei die Normwertsetzung.

### Minimalnormen

Um eine Person als „gesund“ oder „krank“ zu charakterisieren, ist ein Grenzwert erforderlich. So entscheiden in der ärztlichen Praxis Grenzwerte verschiedener Körperparameter (Blutdruck, -fette, -zucker) über mögliche Interventionen. Um eine Zuteilung von Patienten zu verschiedenen Therapieformen vornehmen zu können, sieht Wydra (2006b,

S. 225) insbesondere die Minimalnorm bei motorischen Messverfahren als bedeutsam.

### *Idealnomen*

Im Gesundheitsbereich werden weiterhin häufig Ideale formuliert. Zum einen wird beispielsweise vom idealen Körpergewicht gesprochen. Zum anderen existieren Angaben zu körperlicher Aktivität und körperlichem Training, die zur Verbesserung der Fitness als ideal angesehen werden (ACSM, Garber et al., 2011). Israel (1990, S. 12) beschreibt diese Norm auch als „Optimalnorm“, die auf den Gesundheitsbereich bezogen, eine hohe „Gesundheitsstabilität, vielseitige Leistungsfähigkeit, Belastbarkeit, Resistenz, Erholungs- und Entspannungsfähigkeit, Wohlbefinden“ beinhaltet. Die Festlegung einer solchen Norm sieht er aufgrund des anzustrebenden Zustandes insbesondere für die Therapie als bedeutsam. Wird die angestrebte Idealnomen auf einen Parameter reduziert, so beschreibt die WHO die „ideale“ Gesundheit mehrperspektivisch als „a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity“ (WHO, 2006, S. 1).

### *Spezialnormen*

Diese Normen kommen unter anderem in der Talentdiagnostik im Leistungssport zur Anwendung, um die Tauglichkeit einer Person für eine bestimmte Disziplin bzw. Sportart zu bewerten (Israel, 1983, S. 362).

## **Standards**

Im Bildungssektor werden die von der Kultusministerkonferenz formulierten Bildungsstandards in Mindest-/Minimal-, Regel- und Exzellenz-/Maximalstandards unterschieden (KMK, 2005, S. 9), die sich auch auf den Gesundheitsbereich übertragen lassen.

### *Mindest-/Minimalstandards*

Zu Beginn einer Rehabilitation kann eine Mindestkompetenz formuliert werden, die ein Patient in einem Motoriktest bzw. -screening erreichen soll, um an bestimmten Therapiegruppen teilnehmen zu können. Dies gilt auch für das Ende einer beispielsweise zweiwöchigen Rehabilitationsphase. Wird eine Mindestnorm unterschritten, können weitere ambulante Maßnahmen angeordnet werden.

### *Regelstandards*

Diese Standards können bei einem motorischen Screening zum Einsatz kommen, um beispielsweise den durchschnittlichen Zustand von Patienten bestimmter Krankheitsbilder nach einer zweiwöchigen Anschlussheilbehandlung zu charakterisieren.

### *Exzellenz-/Maximalstandards*

Weiterhin können Testnormwerte definiert werden. Personen, die über einem Testwert liegen, können an Gesundheitssportgruppen teilnehmen.

### **Diagnostische Güte**

Grenzwertbestimmungen durch Mittelwerte geben keine Antwort auf spezifische Fragestellungen. Dies ist erst durch die Bestimmung der diagnostischen Güte möglich (Abb. 10, S. 34). Dabei können je nach Fragestellung unterschiedliche Cut-Off-Werte bestimmt werden, anhand derer die Ergebnisse als negativ oder positiv beurteilt werden. Die rechnerische Bestimmung eines Normwertes für ein motorisches Screening kann hierbei mit der von Israel (1983) definierten Minimalnorm oder mit den im Bildungsbereich gebräuchlichen Mindeststandards verglichen werden.

Neben einer rechnerischen Herangehensweise sollte die Wahl des Cut-Off-Wertes grundsätzlich jedoch auch von dem Praxisnutzen des Screenings abhängig sein. Dieser entscheidet, ob eine höhere Sensitivität oder Spezifität angestrebt werden soll (Lalkhen, & McCluskey, 2008, S. 221).

Auf dem Weg zur Findung eines Normwertes sollte retrospektiv überprüft werden, ob das gewünschte Ergebnis tatsächlich abgebildet wird. Hier zeigte sich in Metaanalysen, dass Urteilsfällungen, die aus statistischen Vergleichen großer Stichproben resultieren, valider sind als Urteile, die subjektiv z. B. durch einen Therapeuten zustande kommen (Ægisdóttir et al., 2006; Grove, Zald, Lebow, Snitz, & Nelson, 2000).

### **ICF-Beurteilungsmerkmale**

Die WHO entwickelte im Rahmen der ICF eine weitere Möglichkeit zur Grenzwertbestimmung. Das Ausmaß eines beispielsweise motorischen Problems (nicht vorhanden bis voll ausgeprägt) soll anhand von Prozentwerten der Testergebnisse bestimmt werden (DIMDI, 2005, S. 27; Tab. 7).

*Tabelle 7: ICF-Grenzwertbestimmung (modifiziert nach DIMDI, 2005, S. 27)*

<b>Ausmaß des Problems</b>	<b>Prozentwerte</b>
Problem nicht vorhanden (ohne, kein, unerheblich, ...)	0 - 4
Problem leicht ausgeprägt (schwach, gering, ...)	5 - 24
Problem mäßig ausgeprägt (mittel, ziemlich, ...)	25 - 49
Problem erheblich ausgeprägt (hoch, äußerst, ...)	50 - 95
Problem voll ausgeprägt (komplett, total, ...)	96 - 100

### 3.4 Zusammenfassung

Das Sammeln und Bewerten personenbezogener Daten zur Planung therapeutischer Interventionen gehört im klinischen Kontext zur täglichen Praxis. Innerhalb eines diagnostischen Prozesses stellen Screeningtests eine ökonomische Methode dar. Werden bei allgemeinen Tests meist mehrere beispielsweise motorische Merkmale einer Person untersucht, sollen Screenings einen ersten Eindruck über ein Merkmal vermitteln und insbesondere das Nebengütekriterium der Ökonomie erfüllen. Beide Methoden unterscheiden sich zudem hinsichtlich der Evaluationserfordernisse. Neben einer Beurteilung der Testgütekriterien sollte auch die diagnostische Güte von Screenings überprüft werden. Hierbei wird ein Trennpunkt (Cut-Off-Wert) bestimmt, unterhalb oder oberhalb dessen ein Merkmal als auffällig (positiv) oder unauffällig (negativ) charakterisiert werden kann. Diese Form der Normwertsetzung erlaubt gerade bei Screenings im Gegensatz zu den im Gesundheitssektor häufig benutzten Normen oder im Bildungsbereich aufgestellten Standards eine rein rechnerische Herangehensweise. Eine weitere Möglichkeit ist die Betrachtung von Ergebnisverteilungen.

Aufgrund der sich durch Bestimmung der diagnostischen Güte ergebenden dichotomen Ergebnisse lassen sich diese Verfahren in sequentielle Diagnosestrategien integrieren. Nach medizinischer Diagnose kann beispielweise die Mobilität als positiv oder negativ beurteilt werden. So sind lediglich Personen mit positiven Ergebnissen einer weiterführenden Diagnostik zuzuführen, um mögliche Einschränkungen zu spezifizieren. Dies optimiert den diagnostischen Prozess. Die Entwicklung und Evaluation solcher Verfahren kann demnach als sinnvoll angesehen werden. Daher wurde im Rahmen dieser Arbeit ein ICF-orientiertes motorisches Screeningverfahren zum Einsatz im Gesundheitsbereich entwickelt.

## 4 Entwicklung eines neuen Screenings

Im Folgenden werden die Schritte zur Entwicklung eines weiterentwickelten Timed Up and Go Tests (TUG; Podsiadlo, & Richardson, 1991) aufgezeigt. Es erfolgt zunächst eine Erläuterung des klassischen Tests. Anschließend wird die Anwendungsbreite anhand von Studien aufgezeigt, Einflussfaktoren, Durchführungsvarianten und Test- bzw. diagnostische Güte diskutiert sowie bisherige Modifikationen erläutert. Im zweiten Schritt werden die Entwicklungsstufen der Testaufgaben des neuen TUG<sub>plus</sub> beschrieben.

### 4.1 Klassischer Timed Up and Go Test (TUG)

Testbatterien wie z. B. der Tinetti-Test (Tinetti et al., 1986) und die Berg-Balance-Scale (Berg et al., 1989) bestehen aus 20 bzw. 14 Items und weisen einen hohen Durchführungsaufwand bzw. eine lange Testzeit auf. Der Einzeltest „Get Up and Go“ (GUG; Mathias, Nayak und Isaacs, 1986) besteht lediglich aus vier Items (von einem Stuhl aufstehen, eine drei Meter entfernte Pylone umrunden und wieder hinsetzen) und sieht ein subjektives Bewertungsschema der Bewegungsqualität vor.

Beim klassischen TUG wird die Mobilität in der Geriatrie ökonomischer erfasst. Hierbei werden die gleichen Items wie beim GUG genutzt, unter Verwendung des objektiven Parameters Zeit (Abb. 14). In Anlehnung an Podsiadlo und Richardson (1991) zeigen Personen, die mehr als 10 s (Cut-Off-Wert) zur Absolvierung der Items benötigen, bei diesem Test Probleme hinsichtlich der Mobilität.

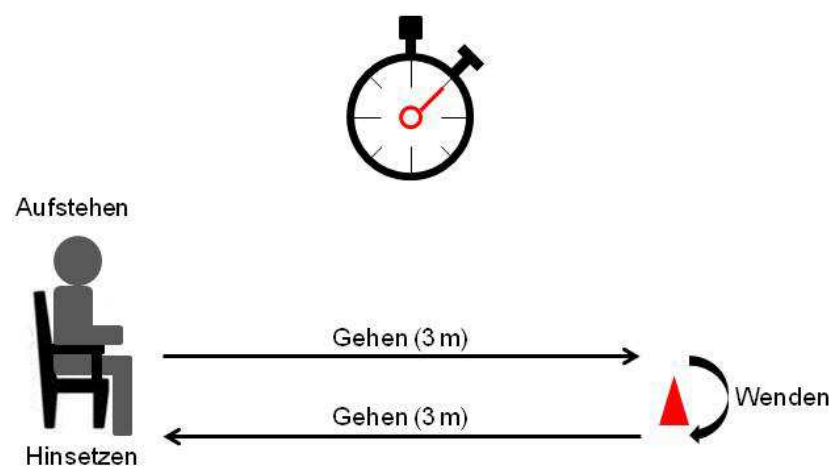


Abbildung 14: Schematischer Aufbau des Timed Up and Go Tests nach Podsiadlo und Richardson (1991)

### 4.1.1 Anwendung

Der Test wird in Untersuchungen bei Personen im Alter von 80 bis 89 Jahren (Corrie, Brooke, Mansfield, Cowley, Morris, & Masud, 2015; Klima, Newton, Keshner, & Davey, 2013; Mulasso, Roppolo, Liubicich, Settanni, & Rabaglietti, 2015; Nordin, Lindelöf, Rosendahl, Jensen, & Lundin-Olsson, 2008) und über 90 Jahren (Cadore et al., 2014) angewendet.

Auch bei jüngeren Altersgruppen zwischen 60 und 69 Jahren (Bird, Hill, & Fell, 2012; De Araújo, De Oliveira, Martins, De Moura Pereira, Copetti, & Safons, 2013; Pertille et al., 2012; Sakurai et al., 2013) und sogar bei 50- bis 59-Jährigen (Sucuoglu, Tuzun, Akbaba, Uludag, & Gokpinar, 2015) kommt der Test zum Einsatz. Zudem zeigte eine Untersuchung, dass dieser auch bereits mit Kindern ab drei Jahren möglich ist (Williams, Carroll, Reddihough, Phillips, & Galea, 2005).

Grundsätzlich kann das Testergebnis von verschiedenen Faktoren beeinflusst werden.

### 4.1.2 Einflussfaktoren

#### *Alter*

Testzeiten des TUG wurden in der Vergangenheit häufig in Abhängigkeit des Alters ermittelt (Tab. 8).

*Tabelle 8: Altersabhängige Testzeiten (s) beim TUG*

Quelle	Alter (Jahre)						
	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80 - 99
Bohannon (2006) - Metaanalyse	-	-	-	-	8,1 s (n=176)	9,2 s (n=798)	11,3 s (n=1102)
Pondal, & del Ser (2008)	-	-	-	-	-	9,5 - 9,9 s (n=225)	11 - 12 s (n=78)
Klima et al. (2013)					10,4 s (n=56)		13,4 s (n=75)
Bird et al. (2012)	-	-	-	-	6,3 s (n=27)	-	-
Tsuji et al. (2014)	-	-	-	-	4,9 s (n=18)	-	-
Vereeck et al. (2008)	4,4 s (n=27)	4,6 s (n=22)	4,9 s (n=8)	5,6 s (n=7)	6,7 s (n=7)	7,8 s (n=20)	-

Bohannon (2006) zeigte im Rahmen einer Metaanalyse anhand von 21 Studien, dass die Gehgeschwindigkeit beim TUG mit zunehmendem Alter abnimmt (s. Tab. 8). Die Testzeiten wurden zu einem späteren Zeitpunkt in einer Studie mit 308 Personen ohne Gehstörung und einem Alter über 71 Jahren von Pondal und del Ser (2008) bestätigt. So nahmen die mittleren Zeiten von 9,5 s bei 71- bis 75-Jährigen bis 12 s bei 86- bis 99-Jährigen zu.

Vereeck, Wuyts, Trujen und Van de Heyning (2008) testeten Personen in der Altersspanne von 25 bis 75 Jahren. Sie konnten zeigen, dass die Zeit des TUG pro Lebensdekade bei einer Korrelation von 0,82 ( $p < 0,001$ ) um bis zu 0,8 s zunahm. Konträr zu diesen Ergebnissen fanden Klima et al. (2013) eine durchschnittliche Zeit von 10,4 s für 60- bis 74-Jährige und 13,4 s für Personen älter als 75 Jahren. In anderen Studien mit Personen im Alter zwischen 60 und 70 Jahren finden sich Testzeiten zwischen 4,9 s (Tsuji, Kitano, Tsunoda, Himori, Okura, & Tanaka, 2014) und 6,3 s (Bird et al., 2012).

### *Geschlecht*

Pondal und del Ser (2008) untersuchten den Einfluss des Geschlechts (171 Männer, 137 Frauen) auf das Ergebnis des TUG. Bei Personen zwischen 71 und 99 Jahren benötigten Männer signifikant weniger Zeit als Frauen. Dies wurde auch in weiteren Studien bestätigt (Vereeck et al., 2008; Samson, Meeuwsen, Crowe, Dessens, Duursma, & Verhaar, 2000; Samson, Crowe, De Vreede, Dessens, Duursma, & Verhaar, 2001). Musselman und Brouwer (2005) konnten bei 70- bis 80-Jährigen hingegen keine geschlechtsspezifischen Effekte feststellen.

### *Körpergewicht*

In einigen Studien konnte ein positiver Zusammenhang zwischen der Testzeit und dem Körpergewicht nachgewiesen werden (Southard, Dave, & Douris, 2010; Vereeck et al., 2008).

### *Kognition*

Die Gehgeschwindigkeit nimmt auch mit zunehmender kognitiver Erkrankung ab. Dies zeigten Vereeck et al. (2008), indem sie die Testzeit (TUG) mit der Global Deterioration Scale – eine Skala zur Beurteilung des Demenzstatus (Reisberg, Cerris, De León, & Crook, 1982), verglichen.

### *Erkrankungen*

Sogar bei Schlaganfallpatienten wird der TUG bevorzugt angewendet, obwohl 80 % per se schon ein auffälliges Gangbild zeigen (Wevers, van de Port, Vermue, Mead, & Kwakkel, 2009). Dementsprechend liegen Personen, bei denen der Schlaganfall immerhin zweieinhalb bis zwölf Monate zurück liegt, mit durchschnittlich 25 bis 28 s immer noch deutlich oberhalb des von Podsiadlo und Richardson (1991) aufgestellten Cut-Off-Wertes von 10 s (s. Kap. 4.1, S. 47; Blennerhassett, & Dite, 2004; Dean, Richards, & Malouin, 2000; Salbach, Mayo, Wood-Dauphinee, Hanley, Richards, & Côte, 2004). Personen, bei denen der Zeitpunkt des Schlaganfalls im Mittel vor 5,4 Jahren war, benötigten lediglich noch durchschnittlich 15 s (Yang, Wang, Lin, Chu, & Chan, 2006).

#### **4.1.3 Störgrößen**

In Studien finden sich Abweichungen hinsichtlich des verwendeten Materials. So werden beispielweise Stühle mit Armlehnen (Asmidawati, Hamid, Hussain, & Hill, 2014; Hoozeboom, Dronkers, Van den Ende, Oosting, & Van Meeteren, 2010) und ohne Armlehnen benutzt (Broekmans, Roelants, Alders, Feys, Thijs, & Eijnde, 2010). Um den Störeffekt (Änderung des Testmaterials) zu untersuchen, testeten Eekhof, De Bock, Schaapveld und Springer (2001), inwiefern die Benutzung unterschiedlicher Stühle das Testergebnis beeinflusst. Hierzu durchliefen 33 ältere Personen (mittleres Alter: 84 Jahre) dreimal den TUG, wobei jedes Mal ein anderer Stuhl benutzt wurde. Die Stühle unterschieden sich hinsichtlich der Sitzhöhe, des Vorhandenseins einer Armlehne, der Sitztiefe, Lehnenneigung und des Sitzpolsters. Es zeigte sich, dass die Testpersonen lediglich bei der Benutzung eines Sessels mehr Zeit (längere Aufstehzeiten) benötigten. Das Vorhandensein einer Armlehne hatte keinen Einfluss auf das Testergebnis.

In Studien unterscheiden sich weiterhin die Anzahl absolvierter Durchgänge. So finden sich Studien, in denen kein (Asmidawati et al., 2014; Schacht, & Ringe, 2012) oder ein Probeversuch (Regterschot, Folkersma, Zhang, Baldus, Stevens, & Zijlstra, 2014) durchgeführt wurde. Hinsichtlich der Wertung des Tests werden entweder ein Durchgang (Asmidawati et al., 2014; Schacht, & Ringe, 2012), zwei Versuche bei Wertung des Schnellsten (Sikjaer, Rolighed, Hess, Fuglsang-Frederiksen, Mosekilde, & Rejnmark, 2014) oder drei Durchgänge absolviert (Maanum, Jahnsen, Frosli, Larsen, & Keller, 2011; Regterschot et al., 2014). Hierbei wird ebenfalls der schnellste Versuch gewertet.



Weiterhin finden sich in der Literatur unterschiedliche Testinstruktionen. So soll der Test entweder schnellstmöglich (Pertille et al., 2012; Sakurai et al., 2013), schnellstmöglich und sicher (Kumar et al., 2015; Maanum et al., 2011), im gewohnten Gehtempo (Gleeson, Sherrington, & Keay, 2014; Granacher, Lacroix, Muehlbauer, Roettger, & Gollhofer, 2013; Lee, Chang, Tsao, Hung, Huang, & Lin, 2013; Nordin et al., 2008) oder in einem selbst gewählten Tempo (Klima et al., 2013; Regterschot et al., 2014) durchgeführt werden. Häufig werden auch keine Angaben zum Tempo gemacht (Gaßner, Janzen, Schwirtz, & Jansen, 2014; Mulasso et al., 2015).

Die sich in Abhängigkeit der Testinstruktion ergebenden Testzeiten sind anhand einiger Studien in Tabelle 9 dargestellt.

*Tabelle 9: Instruktionsabhängige Testzeiten (s) beim TUG*

Instruktion	Schnellstmöglich	Schnellstmöglich und sicher	Gewohntes Gehtempo	Selbstgewähltes Gehtempo	Keine Instruktion
<b>Zeit (s)</b>	<b>5 - 9</b>	<b>6 - 7</b>	<b>9 - 23</b>	<b>12</b>	<b>25</b>
<b>Alter (Jahre)</b>	62 - 84	39 - 55	63 - 80	> 60	83
<b>Stichprobe (Quelle)</b>	Ältere (Pertille et al., 2012; Regterschot et al., 2014)  Übergewichtige (Sakurai et al., 2013)	Personen mit Hüftarthrose (Kumar et al., 2015)  Personen mit Zerebralparese (Maanum et al., 2010)	Ältere (Granacher et al., 2013)  Personen mit Sturzrisiko (Lee et al., 2013) <u>Sturzrisiko:</u> Leicht: 11 - 12 s Moderat: 14 - 15 s Hoch: 20 - 23 s	Ältere (Klima et al., 2013)	Bewohner aus Seniorenheimen (Mulasso et al., 2015)

Es zeigt sich, dass die Art der Anweisung, neben den in Kapitel 4.1.2 (S. 48) aufgeführten Einflussfaktoren, die Testzeit maßgeblich beeinflusst.

Die Anweisung „schnellstmöglich und sicher“, die in den Studien von Kumar et al. (2015) und Maanum et al. (2010) vorzufinden ist (s. Tab.9), stellt hierbei schon ein Widerspruch in sich dar. Ältere Probanden zwischen 62 und 84 Jahren schneiden mit der Instruktion „schnellstmöglich“ genauso gut ab, wie deutlich jüngere Probanden zwischen 38 bis 55 Jahren mit der Instruktion „schnellstmöglich und sicher“ (Pertille et al., 2012).

#### 4.1.4 Testgüte

Bereits Podsiadlo und Richardson (1991) zeigten beim TUG, dass die Ergebnisse unabhängig davon sind, welcher Testleiter den Test durchführt und auswertet. Dies konnte auch in späteren Studien hinsichtlich der Interrater-Reliabilität zu weiteren Personengruppen und Krankheitsbildern bestätigt werden (Schoppen et al., 1999; Shumway-Cook et al., 2000; Morris et al., 2001; Nordin et al., 2006). Zudem zeigte sich, dass die Ergebnisse bei zwei verschiedenen Testzeitpunkten und demselben Beurteiler (Intrarater-Reliabilität) konstant blieben (Schoppen et al., 1999; Nordin et al., 2006). Zur Berechnung wurde entweder Pearsons  $r$  oder der Intraklassenkoeffizient (ICC) verwendet.

Die Zuverlässigkeit (Reliabilität) wurde anhand der Test-Retest-Reliabilität untersucht. Hinsichtlich Schlaganfall- (Flansbjer et al., 2005; Ng, & Hui-Chan, 2005), Parkinson- (Huang et al., 2011) und Alzheimer-Patienten (Ries et al., 2009) ergaben sich hohe Reliabilitäten. Rockwood et al. (2000) konnte in einer Studie mit großer Personenstichprobe, bestehend aus älteren Personen mit und ohne kognitive Probleme, die Reliabilität nicht bestätigen. Die Autoren sehen die Begründung darin, dass eine lange Zeitspanne (durchschnittlich 16 Wochen) zwischen beiden Testzeitpunkten lag und beide Testdurchführungen von unterschiedlichen Beurteilern bewertet wurden (Rockwood et al., 2000, S. M72). Bös (2017) empfiehlt für Retest-Analysen daher ein maximal zweiwöchiges Zeitintervall.

Die Validität des Verfahrens wurde anhand folgender Außenkriterien bewertet: Tinetti-Test (Lin et al., 2004), Berg-Balance-Scale (Bennie et al., 2003; Podsiadlo & Richardson, 1991), Sickness-Impact-Profile (Schoppen et al., 1999), Groningen Activity Restriction Scale (Schoppen et al., 1999), Barthel-Index und Ganggeschwindigkeit (Podsiadlo, & Richardson, 1991). Die Korrelationen waren hinsichtlich des Tinetti-Tests, der Berg-Balance-Scale und des Barthels-Index hoch. Die Beziehungen zum Sickness-Impact-Profile, zur Groningen Activity Restriction Scale und zur Ganggeschwindigkeit waren eher moderat. Aufgrund der signifikanten Korrelationen zu den herangezogenen Testverfahren wird die Validität des TUG insgesamt bestätigt.

Tabelle 10 gibt eine Übersicht zu Studien, die die Testgüte beim TUG erfüllt haben und zeigt die jeweilige Ausprägung der statistischen Analysen.

*Tabelle 10: Studien zur Objektivität, Reliabilität und Validität des TUG (ICC = Intraklassenkorrelation,  $r$  = Korrelationskoeffizient; alle Ergebnisse  $p \leq 0,05$ )*

Kriterium	Studie	Stichprobe (n)	Alter (Jahre; M)	Testgüte
<b>Objektivität</b> (Interrater-Reliabilität)	Podsiadlo, & Richardson, 1991	Ältere (22), 3 Rater	Keine Angaben	ICC = 0,99
	Shumway-Cook et al., 2000	Ältere: Stürzer (15), Nichtstürzer (15), 2 Rater	Stürzer 86, Nichtstürzer 78	ICC = 0,98
	Morris et al., 2001	12 Parkinson-Patienten, 12 gesunde Personen, 6 Rater (3 Kliniker, 3 Nichtkliniker)	69	Kliniker (ICC = 0,99), Nichtkliniker (ICC = 0,87)
	Schoppen et al., 1999	Personen mit einseitiger Beinamputation (32), 2 Rater	73	$r = 0,96$
	Nordin et al., 2006	Personen aus Pflegeheimen (78)	85	ICC = 0,91
<b>Intrarater-Reliabilität</b>	Schoppen et al., 1999	Personen mit einseitiger Beinamputation (32), 1 Rater	73	$r = 0,93$
	Nordin et al., 2006	Personen aus Pflegeheimen (78)	85	ICC = 0,92
<b>Reliabilität</b> (Test-Retest-Reliabilität)	Rockwood et al., 2000	Ältere Personen mit und ohne kognitive Probleme (2305)	78	Personen mit kognitiven Problemen: ICC = 0,56  Personen ohne kognitive Probleme: ICC = 0,50
	Huang et al., 2011	Parkinson-Patienten (72)	68	ICC = 0,80
	Flansbjer et al., 2005	Schlaganfall-Patienten; 6 Monate bis 4 Jahre nach dem Anfall (50)	58	ICC = 0,96
	Ries et al., 2009	Alzheimer-Patienten (51)	81	ICC = 0,99
	Ng, & Hui-Chan, 2005	Gesunde (10), Schlaganfallpatienten (11)	Gesunde (64), Patienten (61)	ICC = 0,95
<b>Validität</b> Kriteriumsvalidität	Lin et al., 2004	Ältere (1200)	73	TUG, Tinetti Test: $r = -0,55$
	Bennie et al., 2003	Personen aus rehabilitativen Programmen (20)	38 - 86	$r = -0,47$ (Berg-Balance-Scale)
	Schoppen et al., 1999	Personen mit einseitiger Beinamputation (32)	73	TUG, Sickness Impact Profile 68: $r = 0,40$  TUG, Groningen Activity Restriction Scale: $r = 0,39$
	Podsiadlo, & Richardson, 1991	Ältere (22),	-	TUG, Barthel-Index: $r = -0,51$  TUG, Berg-Balance-Scale: $r = -0,81$  TUG, Gehtempo: $r = -0,61$

#### 4.1.5 Diagnostische Güte

Die Festlegung von Testnormwerten für den TUG beruht entweder auf den in der Praxis gesammelten Erfahrungen, Mittelwertsberechnungen (Wydra, 2006b) oder auf berechneten Trennpunkten und der Beurteilung der diagnostischen Güte (Hawkins, 2005).

Ersteres wurde bereits bei der Erstveröffentlichung des TUG vorgenommen. Die durchschnittlichen Testzeiten wurden mit den Mobilitätskriterien des Barthel-Index in Beziehung gesetzt. Hieraus resultierte ein Trennpunkt bei zehn Sekunden (Cut-Off-Wert). Folglich waren bei Personen, die mehr als zehn Sekunden zur Absolvierung des TUG benötigten, Mobilitätseinschränkungen zu erwarten (Podsiadlo, & Richardson, 1991, S. 146). Es handelte sich um Personen, die zum Testzeitpunkt keine krankheitsbedingten, motorischen Einschränkungen aufwiesen wie z. B. Parkinson, Multiple-Sklerose oder Schlaganfall.

Die diagnostische Güte wurde in der Vergangenheit häufig bestimmt, um die Vorhersagekraft des Tests für das Sturzrisiko von älteren Personen zu beurteilen. Schoene, Wu, Mikolaizak, Menant, Smith, Delbaere und Lord (2013) berichteten in einem Review und einer Metaanalyse (53 Studien, n = 12832) über die prognostische Validität des TUG im Hinblick auf Stürze bei älteren Personen. Die Cut-Off-Werte variierten zwischen acht und 30 Sekunden.

Eine Auswahl an Studienergebnissen ist in Tabelle 11 dargestellt.

*Tabelle 11: Studien zur diagnostischen Güte des TUG (AUC = Area Under the Curve; SEN = Sensitivität; SPE = Spezifität; LR+/- = positive/negative Likelihood-Ratio; PPV/NPV = positiver/negativer prädiktiver Wert; YI = Youden-Index)*

Quelle	Zielgruppe	Alter (Jahre; M $\pm$ SD)	N	Cut-Off-Wert	Diagnostische Güte
Shumway-Cook et al., 2000	Ältere mit Sturzrisiko	Nicht Stürzende 78 $\pm$ 6	15	13,5 s	SEN = 87 % SPE = 87 %
		Stürzende 86 $\pm$ 6	15		
Rose et al., 2002	Ältere mit Sturzrisiko	Stürzende 78 $\pm$ 7	63	Sturzrisiko 10 s	SEN = 71 % SPE = 89 %
		Nicht Stürzende 77 $\pm$ 6	71		
Bischoff et al., 2003	Vergleich der Wohnverhältnisse (Wohngemeinschaft und Pflegeheim)	Personen in Gemeinschaftswohnungen 73 $\pm$ 3	413	12 s	AUC = 0,97 PPV = 76 % NPV = 98 %
		Personen in Pflegeheimen 80 $\pm$ 4	78		
Chiu et al., 2003	Differenzierung von älteren stürzenden und nicht stürzenden Personen	Stürzende (einmalig) 82 $\pm$ 8	17	Einmaliger Sturz vs. kein Sturz: 20,1 s	SEN = 59 % SPE = 88 % AUC = 0,72
		Nicht Stürzende (gematched) 82 $\pm$ 8	22	Mehrere Stürze vs. keine Stürze: 24,7 s	SEN = 77 % SPE = 88 % AUC = 0,9
		Stürzende (mehrfach) 83 $\pm$ 7		Einmaliger Sturz vs. mehrere Stürze: 23,3 s	SEN = 88 % SPE = 82 % AUC = 0,9
Lin et al., 2004	Personen über 65 Jahre mit Sturzrisiko	73	1200	Sturzrisiko: 13 s	AUC = 0,6
Whitney et al., 2004	Vestibuläre Störungen	60 $\pm$ 17	103	11,2 - 13,5 s	SEN = 47 % SPE = 81 %
Kristensen et al., 2007	Ältere mit Sturzrisiko (6 Monate nach Hüftfraktur Operation)	80	79	24 s	SEN = 95 % SPE = 35 % PPV = 41 % NPV = 93 % LR+ = 1,5 LR- = 0,1
Alexandre et al., 2012	Ältere mit Sturzrisiko	Stürzende: 67 $\pm$ 6	19	Sturzrisiko: 12,5 s	SEN = 74 % SPE = 66 % LR+ = 2,2 LR- = 0,4 PPV = 50 % NPV = 84 % AUC = 0,7
		Nicht Stürzende: 66 $\pm$ 5	41		
Nordin et al., 2008	Personen in Pflegeheimen	85 $\pm$ 6	78	Prognostische Validität von Stürzen: 15 s	SEN = 96 % SPE = 32 % LR+ = 1,4 LR- = 0,1
Haines et al., 2008	Ältere mit Sturzrisiko (Stationäre Geriatrie, Rehabilitation)	76 $\pm$ 14	602	Stürze Inzidenzrate: 30 s Ereignisrate: 25 s	SEN = 76 % SPE = 44 % YI = 0,2
Shimada et al., 2009	Ältere mit Sturzrisiko	81 $\pm$ 7	455	Sturzrisiko 16 s	SEN = 53 % SPE = 63 %
Savva et al., 2012	Identifikation von gebrechlichen Personen	65+	1814	Gebrechlichkeit 16 s	SEN = 29 % SPE = 98 % PPV = 52 %

Kojima et al., 2015	Ältere mit Sturzrisiko	Stürzende 73 ± 6	59	13 s	SEN = 31 % SPE = 90 % PPV = 46 % NPV = 81 % LR+ = 2,9 LR- = 0,8 YI = 0,2 AUC = 0,6
		Nicht Stürzende 73 ± 6	200		

#### 4.1.6 Testmodifikationen

Der Test wurde in der Vergangenheit mehrfach modifiziert. So existieren weitere Versionen, bei denen auf einer festgelegten Strecke die Zeit gemessen wird (Abb. 15).

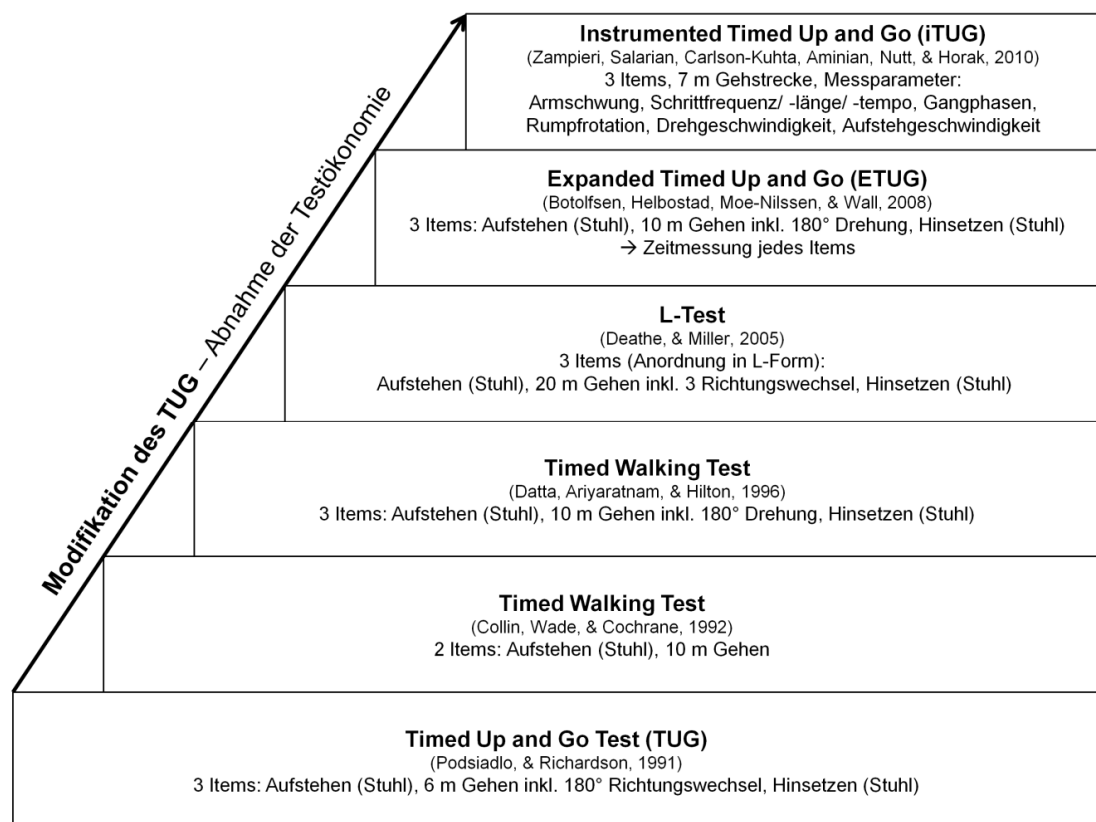


Abbildung 15: Weiterentwicklungen des TUG

Beim Timed Walking Test von Collin, Wade und Cochrane (1992) wurde die Gehstrecke von sechs auf zehn Meter (ohne 180°-Drehung) vergrößert, mit der Begründung die Gehfähigkeit von Personen mit amputierten Gliedmaßen zu testen. Ab einer Zeit von 20 s ist bei diesem Test mit motorischen Auffälligkeiten zu rechnen (Collin et al., 1992, S. 20).

Datta, Ariyaratnam und Hilton (1996) erweiterten diesen Test später durch Integration einer 180°-Drehung. Hierdurch sollte wie beim TUG

die Balance der Testpersonen beurteilt werden. Die Kriteriumsvalidität wurde von Datta et al. (1996) bestätigt.

Aufgrund der eher geringen Eignung als Evaluationsinstrument einer Therapie, wurde der Test später für diese Patientengruppe zum L-Test modifiziert (Deathe, & Miller, 2005). Hierzu wurde die Gehstrecke erneut erweitert (20 Meter Strecke) und in L-Form angeordnet. Zusätzlich sollten nun drei Richtungswechsel durchgeführt werden.

Der Expanded Timed Up and Go (ETUG; Botolfsen, Helbostad, Moenilssen, & Wall, 2008) ist eine Modifikation zur Erhöhung des Informationsgehalts für die Therapie. Hierfür wurde eine separate Zeitmessung jedes Items als zusätzlicher Messparameter herangezogen (95%-Konfidenzintervall: Sitzen - Stehen 0,6 - 0,9 s; Gehen bevorzugtes Tempo 0,5 - 1,0 s; 180°-Richtungswechsel 0,7 - 1,0 s; Gehen schnelles Tempo 0,5 - 1,0 s; Hinsetzen 0,8 - 1,0 s). Der ETUG wies hinsichtlich der Gesamt- und Itemteilzeiten mittlere bis hohe Intra-, Interrater- und Test-Retest-Reliabilitäten ( $ICC = 0,54 - 0,97$ ) sowie eine interne Konsistenz von Cronbachs Alpha  $\alpha = 0,74$  auf (Botolfsen et al., 2008, S. 99).

Zur Erhöhung des Informationsgewinns insbesondere für die Therapie von Parkinsonpatienten wurden später beim sogenannten Instrumented Timed Up and Go Test (iTUG) neben der Zeit weitere Messparameter mittels Körpersensoren erhoben (Zampieri, Salarian, Carlson-Kuhta, Aminian, Nutt, & Horak, 2010). Auf einer sieben Meter Gehstrecke dienten die Aufstehgeschwindigkeit, der Armschwung, die Schrittlänge und -frequenz, das Schrittempo, die Gangphasen, die Rotationsbewegungen und Drehgeschwindigkeit zur Analyse. Diese Parameter wurden hinsichtlich ihrer diagnostischen Güte bzw. ihrer Aussagekraft auf Unterscheidung zwischen Parkinsonpatienten und gesunden Personen untersucht. Zur Unterscheidung von jeweils zwölf Parkinsonpatienten und „Gesunden“ zeigte sich, dass das Armschwungtempo ( $AUC = 0,96$ ) und der Armschwungbewegungsumfang ( $AUC = 0,91$ ) geeignete Kriterien darstellen (Zampieri et al., 2010, S. 173).

#### **4.1.7 Zusammenfassung**

In der Geriatrie und bei Menschen mit Behinderung ist der TUG seit den 1990er Jahren ein häufig verwendetes Screening zur Beurteilung der Motorik und Mobilität. Insbesondere aufgrund der hohen Testökonomie hat sich dieser in der klinischen Praxis international bewährt. Um das Verfahren bei unterschiedlichen Krankheitsbildern anzuwenden, entstanden mehrere Testmodifikationen durch Veränderung des Testaufbaus, -materials, der Anzahl der Durchgänge und Instruktion.

Beim klassischen TUG wurde in zahlreichen Studien die Testgüte bestätigt. Ausgehend von der in der Erstveröffentlichung anhand von Mittelwertberechnungen festgelegten 10-s-Schwelle, wurde der Trennpunkt (Cut-Off-Werte) in Abhängigkeit der Fragestellung, jedoch insbesondere zur Identifizierung des Sturzrisikos von älteren Personen, immer wieder angepasst (s. Tab. 11, S. 55).

Zur Herstellung eines größeren Alltagsbezuges sollte der TUG im Folgenden weiterentwickelt werden.



## 4.2 Erweiterter Timed Up and Go Test (TUG<sub>plus</sub>)

### 4.2.1 Entwicklung ICF-basierter Items

Ziel der Testentwicklung war die Erweiterung des klassischen TUG durch Erhöhung des Alltagsbezugs in zusätzlichen Items ohne die Testökonomie zu vernachlässigen. Störquellen, die durch Integration zusätzlicher Items entstehen können, wurden zugunsten der größeren Alltagsnähe in Kauf genommen. Die Orientierung an der Klassifikation der WHO und deren durch die Mobilität vordefinierten Alltagsbezug (d410 - d489; s. Abb. 3, S. 17) ermöglicht eine eindeutige Zuordnung neuer Bewegungsaufgaben nach internationalem Sprachgebrauch.

Zur Entwicklung von Testaufgaben wurde zunächst eine Voruntersuchung durchgeführt (Kaczmarek, Schwarz, & Wydra, 2014a; 2014b). Hierzu durchliefen 115 Patienten (66 ♀, 49 ♂;  $52 \pm 8$  Jahre; Innere Medizin, n = 11; Hals-Nasen-Ohren-Heilkunde, n = 72; Neurologie, n = 32) 23 Einzeltests, die die alltagsmotorische Leistungsfähigkeit erfassen sollten. Die Einzeltests waren den Subkategorien „Gehen“ (5 Items), „Aufstehen“ (5 Items), „Stehen“ (4 Items), „Springen“ (2 Items) und „Gegenstände bewegen“ (7 Items) der ICF-Mobilitätskategorien zugeordnet. Die bestandenen Einzeltests wurden zu Teilsummenscores addiert.

Zur Überprüfung der Itemschwierigkeit der einzelnen Testaufgaben wurde die Normalverteilung der Teilsummenscores mittels Kolmogorov-Smirnov-Test berechnet. Zudem wurden die Lösungshäufigkeiten bestimmt.

Die Summenscores wiesen keine Normalverteilung auf. 76 % der Probanden lösten die einzelnen Items problemlos (Gehen 85 %, Aufstehen 69 %, Stehen 42 %, Springen 89 %, Gegenstände bewegen 93 %). Somit konnte die Itemschwierigkeit der einzelnen Kategorien weitgehend als zu gering eingestuft werden.

Nach der Voruntersuchung und nach Beurteilung der Alltagsrelevanz, Standardisierbarkeit und des Gefahrenpotentials wurden aus den 23 Aufgaben fünf zentrale Items mit typischem Alltagscharakter ausgewählt: Aufstehen, Übersteigen, Gegenstand mit Fuß bewegen, Gegenstand tragen, Hinsetzen.

#### 4.2.2 Ablauf

Für die Itemauswahl war die Praktikabilität entscheidend. Daher wurde Material verwendet, das in herkömmlichen Sporthallen zur Verfügung steht.

Ausgehend vom klassischen TUG wurde die Streckenlänge von drei Meter auf 15 Meter vergrößert. Beim neuen TUG<sub>plus</sub> wird eine Turnbank als Start- und Endpunkt benutzt. Die Items sind in einem Parcours aufgebaut und sollen, wie beim TUG, schnellstmöglich durchlaufen werden. Hierbei wird die Gesamtzeit gemessen, die die Testperson zur Absolvierung aller Items benötigt.

Für die Testinstruktion „schnellstmöglich“ wurde das allseits bekannte Startkommando „Auf die Plätze, fertig, los“ gewählt, was gleichzeitig einen hohen Anreiz impliziert. Während des Tests erfolgt keine zusätzliche Animation, da der Anreiz durch die Erfüllung der Bewegungsaufgaben ausreichend hoch ist und die Durchführung in der Gruppe, sowie damit verbundene Vergleiche mit anderen Gruppenmitgliedern, zusätzlich die Motivation fördert.

Die durch den TUG<sub>plus</sub> abgebildeten ICF-Mobilitätsdomänen sind in Abbildung 16 verdeutlicht.

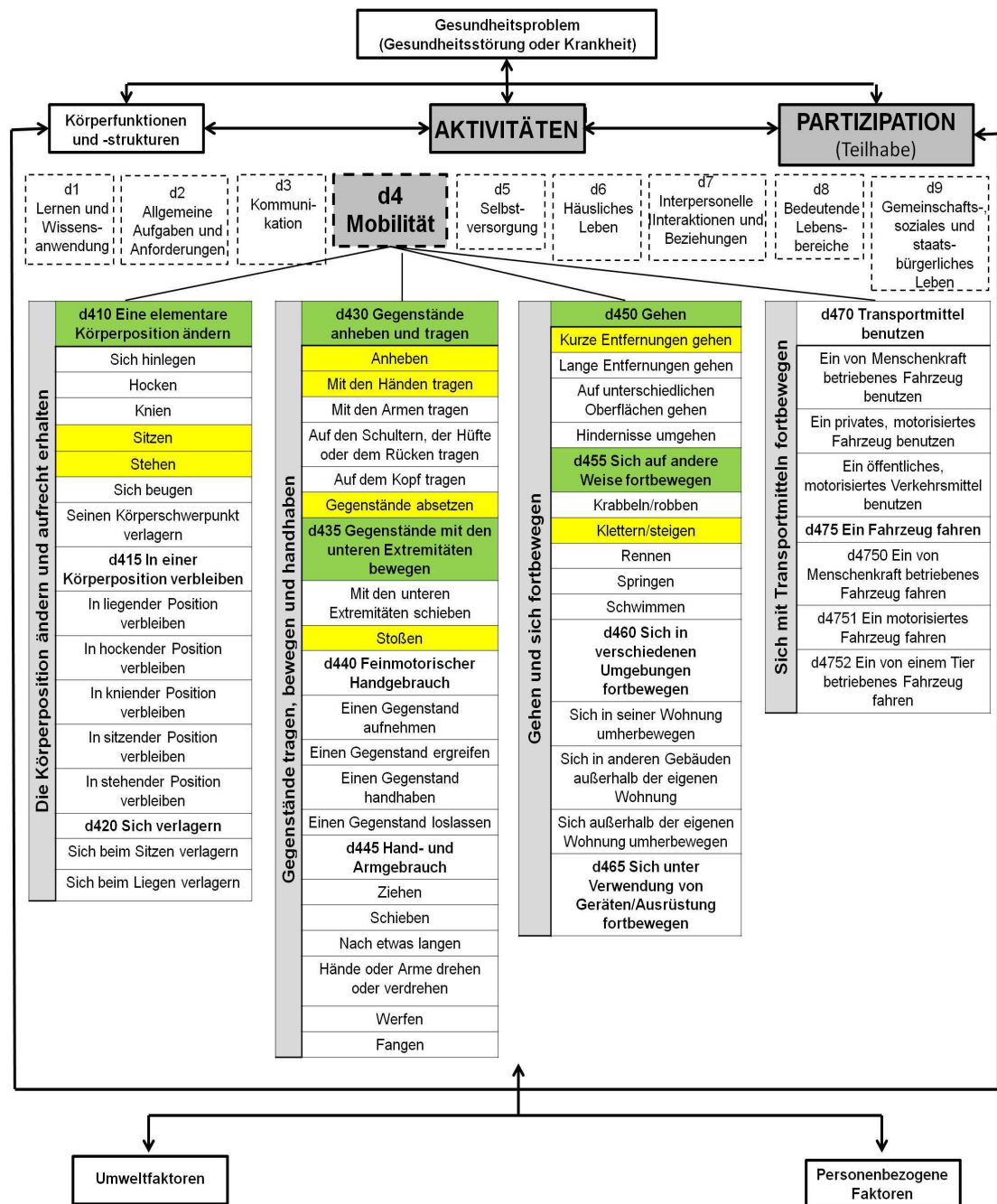


Abbildung 16: Mobilitätsdomänen des TUG<sub>plus</sub> (Haupt- und Unterkategorien farblich markiert)

Das Screening besteht aus den in Abbildung 17 dargestellten Items bzw. motorischen Aufgaben, die als Parcours angeordnet sind.

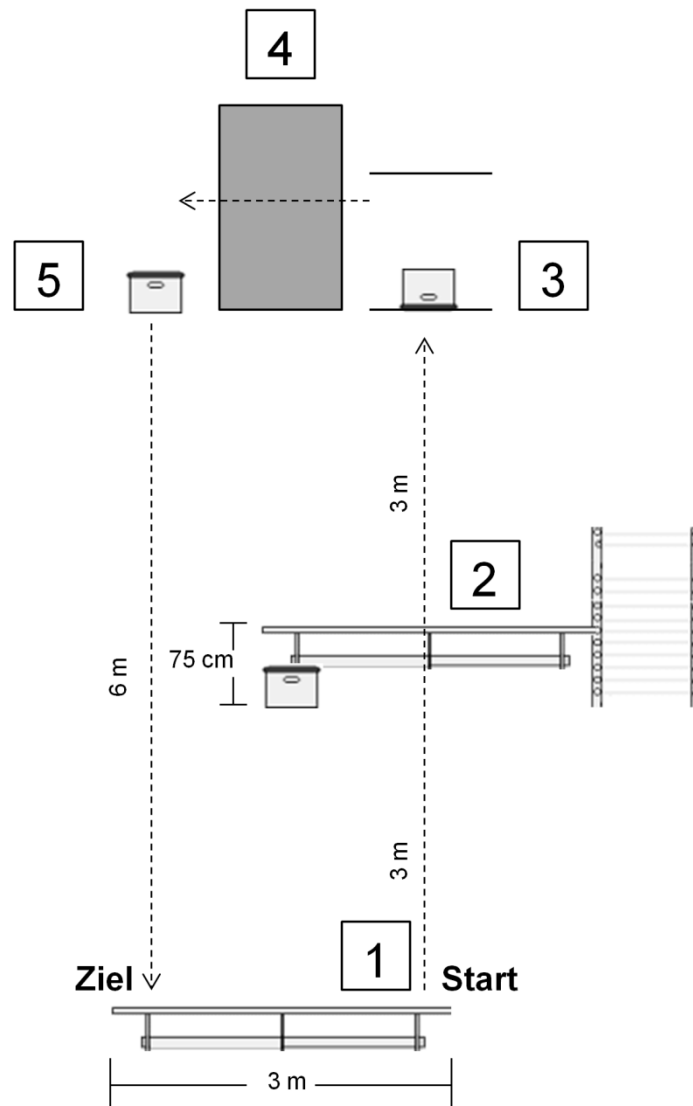


Abbildung 17: Testaufbau und -ablauf des TUG<sub>plus</sub> (1 = Aufstehen (Bank) und Gehen; 2 = Hindernis überwinden; 3 = Kasten umstoßen; 4 = Hinsetzen/Aufstehen (Boden); 5 = Kasten tragen und Hinsetzen (Bank))

*Item 1: Aufstehen (Bank) und Gehen*

Die Testperson sitzt auf einer ca. 35 cm hohen Turnbank. Auf ein Startkommando soll die Person aufstehen und losgehen. Es ist freigestellt, ob beim Aufstehvorgang die Hände benutzt werden oder nicht. Die Zeitmessung beginnt mit dem Startkommando.

*Item 2: Hindernis überwinden*

Nach einer drei Meter Gehstrecke wird ein 75 cm hohes Hindernis überwunden. Dabei kann das Hindernis durch Unterstützung der Hände seitlich überhockt oder sitzend überquert werden.

*Item 3: Kasten umstoßen*

Nach einer erneuten drei Meter Gehstrecke wird ein Turnkasten (18 kg) um 180° mit den Füßen umgestoßen.

*Item 4: Hinsetzen/Aufstehen (Boden)*

Die Testperson setzt sich auf den Boden und soll wieder aufstehen. Beim Sitzen reicht ein kurzer Bodenkontakt.

*Item 5: Kasten tragen und Hinsetzen (Bank)*

Dann hat die Testperson die Aufgabe ein insgesamt 18 kg schweres Gewicht sechs Meter weit zu tragen und dieses vor einer Bank abzustellen. Je nach Kräfteverhältnis des Probanden stehen hierbei folgende Optionen zur Auswahl:

- a) Wegstrecke wird einmal bewältigt: 18 kg schwerer Turnkasten oder 12 kg und 6 kg schwere Getränkekiste tragen.
- b) Wegstrecke wird zweimal bewältigt: 12 kg Getränkekiste sechs Meter tragen, zurück gehen und die restlichen 6 kg tragen.
- c) Die Wegstrecke wird dreimal bewältigt: jeweils eine 6 kg schwere Getränkekiste tragen.

Der Parcours ist beendet, wenn die Testperson wieder auf der Bank sitzt.

Der gegenüber dem TUG höhere motorische Anspruch erfordert im Rahmen eines Vorversuchs das Üben der einzelnen Items. Sind einzelne Items von Personen, z. B. mit höhergradigen motorischen Einschränkungen, nicht zu bewältigen (u. a. Schlaganfall, Multiple Sklerose, Parkinson), kann dies hier auch zum Ausschluss für den eigentlichen Test dienen.

Die Ausführungsalternativen (insbesondere bei Item 5) wurden gewählt, um das Screening für heterogene Gruppen durchführbar zu machen.

Dies war unter anderem die Empfehlung von 16 Experten (sechs Sporttherapeuten, ein Physiotherapeut, neun Sportwissenschaftler; durchschnittliches Alter: 42 Jahre), die in einer Voruntersuchung die Eignung des TUG<sub>plus</sub> bewerten sollten<sup>5</sup>. Weiterhin äußerten sie verschiedene Vor- und Nachteile (Tab. 12).

*Tabelle 12: Expertenbeurteilung: Vor- und Nachteile der Testaufgaben beim TUG<sub>plus</sub>*

<b>Vorteile</b>	<b>Nachteile</b>
Beinhaltet viele alltagsorientierte Bewegungen	Zeitliche Engpässe innerhalb des Therapieplans erschweren die Durchführung (insbesondere bei großen Gruppen)
Geringer Materialaufwand	Standardisierung der Items könnte zu Problemen führen
Schnell durchführbar	Problem der Testauswertung bei Personen, die mindestens ein Item nicht ausführen können
Testung des Ist-Zustandes und zur Verlaufskontrolle	Keine Individualisierung der Hindernishöhe
Eignet sich für fitte Ältere	Wie bei allen Tests wird nur ein Ausschnitt des Lebensalltags abgebildet
Testung mehrerer motorischer Grundeigenschaften im Vergleich zum klassischen TUG	Großer Platzbedarf notwendig
Motorische Defizite können schon bei Personen mittleren Alters aufgedeckt werden	Zu schwierig für körperlich beeinträchtigte Personen
Sinnvoll für Personen bis 70 Jahre	Zu tragender Kasten zu schwer

#### 4.2.3 Bewertung

In einer weiteren Voruntersuchung (Kaczmarek, Fröhlich, Schwarz, & Wydra, 2015) wurde überprüft, ob eine quantitative oder qualitative Bewertung differenzierte Ergebnisse liefert.

An der Studie nahmen insgesamt 123 Patienten mit neurologischen (n = 52) und Hals-Nasen-Ohrenerkrankungen (HNO, n = 71) teil (Alter  $52 \pm 9$  Jahre). Die Patienten absolvierten den TUG<sub>plus</sub>.

Als Testkriterien dienten die benötigte Zeit zur Absolvierung der Items sowie die Bewertung des Parcoursdurchlaufes anhand von qualitativen Kriterien. Zur Aufstellung qualitativer Kriterien wurden im Vorfeld Videoaufnahmen zu Testdurchläufen unterschiedlicher Personengruppen gesichtet.

<sup>5</sup> Die Daten wurden im Rahmen einer im Arbeitsbereich durchgeführten Bachelorarbeit in Zusammenarbeit mit Frau Lisa Marie Paquet erhoben.

Dabei erhielten die Testpersonen bewusst keine Instruktion darüber, wie sie die Items absolvieren sollten. Daraus resultierten unterschiedliche Bewegungsmuster und Lösungsstrategien der jeweiligen Items, die zur besseren Bewertung aufgesplittet und einzeln bewertet wurden (Tab. 13).

*Tabelle 13: Entwicklung eines Beurteilungsbogens*

Testaufgabe	Interpretation
Aufstehen (Bank)	0 = alleine nicht möglich 1 = Abstützen mit Händen 2 = ohne Abstützen der Hände gelöst
Hindernis überwinden	0 = alleine nicht möglich 1 = Bein wird mit Hand über die Bank gezogen 2 = Bank sitzend überwinden 3 = Bank seitlich überhocken
Kasten umstoßen	0 = alleine nicht möglich 1 = > 2 Kastenberührungen 2 = 2 Kastenberührungen 3 = 1 Kastenberührung
Hinsetzen (Boden)	0 = alleine nicht möglich 1 = Person landet auf dem Rücken 2 = ohne Probleme gelöst
Aufstehen (Boden)	0 = alleine nicht möglich 1 = ohne Probleme gelöst
Kasten tragen	0 = kein Gewicht möglich 1 = 3 x 6 kg 2 = 2 x 6 kg + 6 kg 3 = 18 kg Turnkasten oder 12 + 6 kg Getränkekiste
Hinsetzen (Bank)	0 = alleine nicht möglich 1 = Abstützen mit Händen 2 = ohne Abstützen der Hände gelöst

Die gemittelte Zeit der Gesamtstichprobe lag bei 23,2 Sekunden. HNO-Patienten ( $19,7 \pm 5,1$  s) durchliefen den Parcours schneller als neurologische Patienten ( $27,6 \pm 9,1$  s;  $p \leq 0,05$ ).

Die mittlere Aufgabenschwierigkeit ( $p_i$ ) betrug 0,77 (Lösungshäufigkeit:  $77 \pm 14$  %). Die motorischen Aufgaben 2, 3 und 5 lagen im mittleren Schwierigkeitsbereich ( $p_i \geq 0,20 \leq 0,80$ ). Die Aufgaben 1, 4 wiesen eine geringere Schwierigkeit auf ( $p_i > 0,80$ ). Zwischen den Krankheitsbildern bestanden keine Unterschiede hinsichtlich der Aufgabenschwierigkeit.

Für die Testitems 2, 3, 5 waren daher moderate Trennschärfen zu erwarten, da diese im mittleren Schwierigkeitsbereich lagen. Aufgrund der geringen Schwierigkeit der Aufgaben 1 und 4, wurde darüber nachgedacht den Bewertungsmaßstab anzupassen, die Testübungen zu verändern oder die Testaufgaben als Differenzierung für Extremgruppen im Itempool weiterhin zu integrieren.

Welcher Messparameter (Zeit oder Bewegungsqualität) letztendlich verwendet werden soll, wird im Kapitel Evaluation überprüft. Es wird bereits hier vermutet, dass der Parcours nur bei maximal ökonomischer Ausführung (ausreichende Lösung der Bewegungsaufgaben mit möglichst wenig Versuchen) schnellstmöglich durchlaufen werden kann.

Zur Ökonomisierung der Beurteilung der Bewegungsqualität erfolgte eine inhaltliche Reduzierung. Nach Sichtung des Videomaterials wurde deutlich, dass eine schnelle Absolvierung des Parcours ein Abstützen auf der Bank erfordert. So wurde die Bewertung der Aufgaben „Aufstehen, Hinsetzen (Bank)“ entfernt. Weiterhin wurden die Aufgaben „Hinsetzen, Aufstehen (Boden)“ zusammengefasst. Dieses Schema diene für alle weiteren Berechnungen zur Evaluation des TUG<sub>plus</sub> (Tab. 14).

*Tabelle 14: Überarbeiteter Beurteilungsbogen zur Bewertung der Testaufgaben anhand der Bewegungsqualität*

Testaufgabe	Interpretation
Hindernis überwinden	0 = alleine nicht möglich 1 = Bein wird mit Hand über die Bank gezogen 2 = Bank sitzend überwinden 3 = Bank seitlich überhocken
Kasten umstoßen	0 = alleine nicht möglich 1 = > 2 Kastenberührungen 2 = 2 Kastenberührungen 3 = 1 Kastenberührung
Hinsetzen/ Aufstehen (Boden)	0 = alleine nicht möglich 1 = Person landet auf dem Rücken 2 = ohne Probleme gelöst
Kasten tragen	0 = kein Gewicht möglich 1 = 3 x 6 kg 2 = 2 x 6 kg + 6 kg 3 = 18 kg Turnkasten oder 12 + 6 kg Getränkekiste



### 4.3 Zusammenfassung

Die Durchführung diagnostischer Maßnahmen in der Bewegungstherapie kann aus evidenzbasierter Sicht als sinnvoll erachtet werden. Zudem erhöht dies die Legitimation der Disziplin innerhalb der Rehabilitation.

Hinsichtlich der Ausrichtung bewegungsbezogener Diagnostik existieren verschiedene theoretische Ansätze (s. Kap. 2.3, S. 19). Wurden in der Vergangenheit insbesondere Testverfahren konzipiert, die sich am Modell motorischer Fähigkeiten orientieren, erfahren aktuell interdisziplinär ausgerichtete Konzepte größere Akzeptanz. So hat die WHO 2001 mit der ICF eine Klassifikation entwickelt, die für eine einheitliche Sprache in Medizin und Therapie sorgen kann. Zukünftig sollen sich rehabilitative Maßnahmen vermehrt an diesem Modell orientieren.

Die ICF wird dem in mehreren Wissenschaftsdisziplinen aktuell gebräuchlichen Begriff der „Kompetenzen“ gerecht. So können die in dieser Klassifikation unter dem Konstrukt „Mobilität“ formulierten motorischen Fertigkeiten auch als „motorische Qualifikationen“ interpretiert werden. Als Mobilität können somit Fertigkeiten zusammengefasst werden, die Voraussetzung für das Interagieren mit der Umwelt und somit für ein unabhängiges Leben bedeutsam sind.

Zur Diagnostik von Mobilitätseinschränkungen werden vor allem im Gesundheitssport ökonomische Messmethoden favorisiert. Screenings sind hierbei einfach durchzuführen und erlauben auch große Stichproben ökonomisch zu testen. Der TUG ist hierbei ein in der Praxis international etabliertes Screening zur Beurteilung der Mobilität, der in der Vergangenheit mehrfach modifiziert wurde.

Um die Alltagsnähe zu erhöhen, sollte der TUG im Rahmen dieser Arbeit durch Integration weiterer alltagsorientierter Bewegungshandlungen für die Anwendung mit verschiedenen Personen- und Altersgruppen erweitert werden. Der TUG<sub>plus</sub> besteht aus Aufgaben, die den ICF-Mobilitätskategorien „Körperposition ändern und aufrecht erhalten“, „Gehen und sich fortbewegen“ und „Gegenstände tragen und handhaben“ zugeordnet werden können. Dieses Screening soll wie der klassische TUG ökonomische Testungen ermöglichen.

Im Folgenden wird dieses Verfahren hinsichtlich Anwendbarkeit, Testgütekriterien und Grenzwertbestimmung evaluiert.

## 5 Evaluation des TUG<sub>plus</sub>

Das Kapitel beinhaltet die durchgeführten Untersuchungen zur Evaluation des TUG<sub>plus</sub>. Zunächst werden übergeordnete Fragestellungen formuliert und die Untersuchungen inklusive Personenstichprobe sowie verwendete Test- und Statistikverfahren im Überblick dargestellt. Anschließend folgt eine Voruntersuchung zum klassischen TUG und Untersuchungen zum TUG<sub>plus</sub> (Gesamtstichprobe, Reliabilität, Validität und Grenzwertbestimmung).

### 5.1 Übergeordnete Fragestellungen und Methodik

Zur Evaluation des TUG<sub>plus</sub> sollen folgende Fragen beantwortet werden:

- Wie schnell können unterschiedliche Personen- bzw. Patienten- und Altersgruppen den klassischen TUG absolvieren?
- Wie anspruchsvoll ist der TUG<sub>plus</sub> für verschiedene Personen- und Patientengruppen?
- Wie schnell können unterschiedliche Personen- und Patientengruppen den Parcours absolvieren?
- Erfüllt das neue Screening die Gütekriterien?
- Wie ist die diagnostische Qualität zu bewerten?
- Kann ein Trennpunkt für motorisch auffällige und nicht auffällige Personen bestimmt werden?

#### 5.1.1 Untersuchungen

Zur Beantwortung der Fragestellungen wurden mehrere Untersuchungen durchgeführt (Abb. 18). In einer Vor- bzw. Paralleluntersuchung wurden die Testzeiten des klassischen TUG überprüft. Zur Evaluation des TUG<sub>plus</sub> wurden zunächst die Testgütekriterien Reliabilität, Responsivität und Validität (Kriterium, Konstrukt) überprüft. Anschließend sollten Trennpunkte für auffällige bzw. nicht auffällige Testzeiten zunächst anhand deskriptiver Maße (Majoritäts- und Minimalnormen) und anschließend mittels diagnostischer Güte anhand drei Gold- bzw. Referenzstandards (Motorische Leistungsfähigkeit, Gesundheitsstatus, Kombination beider Standards) bestimmt werden.

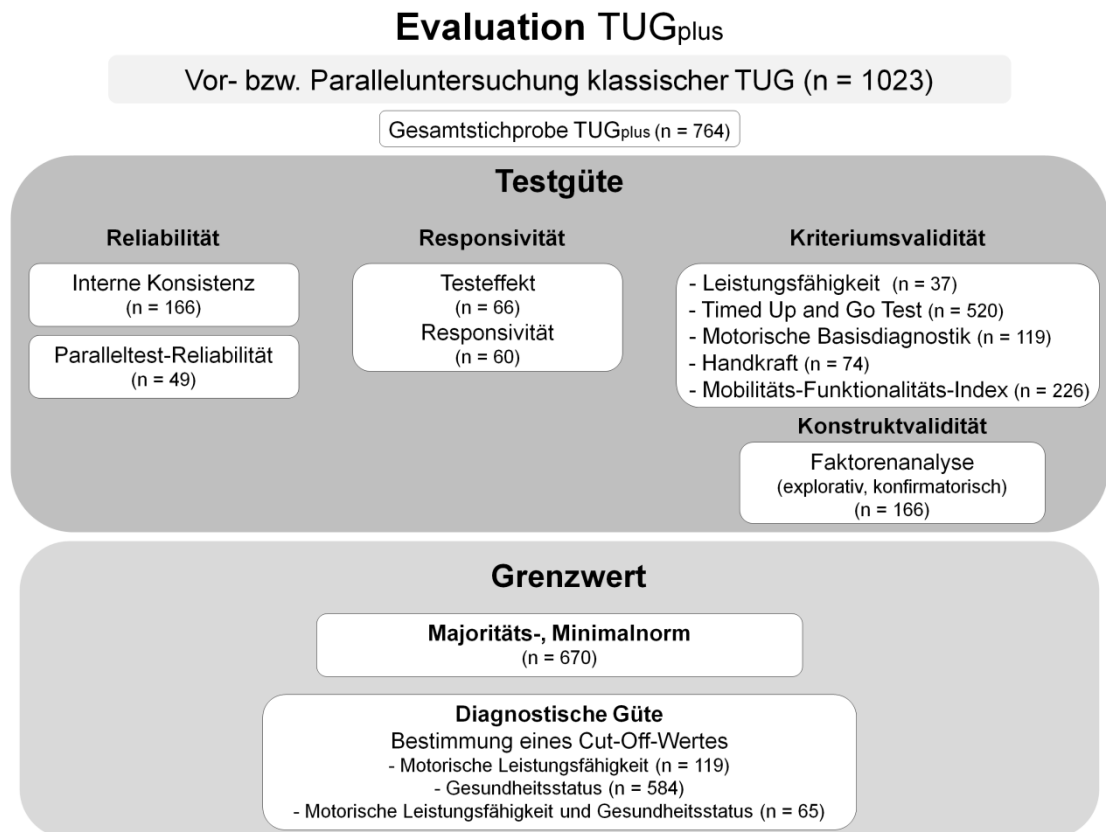


Abbildung 18: Untersuchungen zur Testevaluation (Vor- bzw. Paralleluntersuchung zum klassischen TUG; TUG<sub>plus</sub> – Testgüte, Grenzwertbestimmung) inklusive entsprechenden Stichprobengrößen (n)

### 5.1.2 Stichproben

An den Untersuchungen zum TUG und TUG<sub>plus</sub> nahmen Personen mit verschiedenen Krankheitsbildern und gesunde Probanden teil.

- **Neurologie**

Alle Versuchspersonen absolvierten eine mindestens dreiwöchige Anschlussheilbehandlung in der Bosenberg-Klinik in St. Wendel (Saarland).

Die Patienten litten an unterschiedlichen neurologischen Erkrankungen: Multiple Sklerose, Schlaganfall, Hirnblutung, Schädelhirntrauma, Kleinhirnfarkt, Guillain-Barré-Syndrom, Hirnhautentzündung, Nervenschädigungen, Epilepsie, Myasthenia Gravis, Hirntumor, Aneurysma, Polyneuropathie, ZNS-Lymphom, Spinale Muskelatrophie, Zerebrale arteriovenöse Malformation.

- **Hals-Nasen-Ohrenheilkunde (HNO)**

Auch alle HNO-Patienten absolvierten eine mindestens dreiwöchige Anschlussheilbehandlung in der Bosenberg-Klinik. Die Patienten litten an einer oder mehreren der folgenden Erkrankungen: Tinnitus, Schwindel, Hörsturz, Hyperakusis, Morbus Menière, Taubheit, Neuritis vestibularis, Hörminderung, Gleichgewichtsstörung, Akustikusneurinom.

- **Innere Erkrankungen**

Die Probanden nahmen einmal wöchentlich an Rehabilitationssportgruppen aus verschiedenen Vereinen oder Institutionen im Saarland teil. Sie litten an unterschiedlichen Herz- oder Krebserkrankungen.

- **Orthopädie**

Die Orthopädiepatienten litten an mindestens einer der folgenden Erkrankungen und/oder Verletzungen: Bandscheibenvorfall, Kniearthrose, Totalendoprothese (Hüfte).

- **„Körperlich Gesunde“**

Als „körperlich gesund“ wurden Personen charakterisiert, die zum Testzeitpunkt unter keinen manifesten körperlichen Erkrankungen und/oder Verletzungen litten.

- **Sportstudierende**

Die Sportstudierenden absolvierten zum Testzeitpunkt ein Sportstudium an der Universität des Saarlandes.

Der Ablauf der Untersuchungen orientierte sich an den Richtlinien zur Durchführung von Testverfahren des American College of Sports Medicine (ACSM; Thompson, Arena, Riebe, & Pescatello, 2013). Um Personen mit möglichen Kontraindikationen auszuschließen bzw. den Gesundheitszustand im Vorhinein abschätzen zu können, wurden die Probanden anamnestisch standardmäßig befragt. So konnten die Erkrankungen und Verletzungen durch Selbstangaben und auf der Grundlage medizinischer Befunde spezifiziert werden.

Die Probanden wurden wie folgt rekrutiert:

- **Neurologie, HNO**  
Die Patienten wurden von der Verwaltung der Bosenberg Klinik rekrutiert. Voraussetzung zur Teilnahme war die ärztliche Freigabe. Zur Kontrolle der Sekundärvarianz wurde die Erfahrung der Personen insofern konstant gehalten, dass jeder Patient in der ersten Rehabilitationswoche getestet wurde. Neurologische Patienten wurden zwei Tage und HNO-Patienten drei Tage nach Ankunft in der Klinik getestet. Um den Einfluss verschiedener Parameter (sportlicher Aktivität bzw. sportlicher Vergangenheit, Medikamentenkonsum, Voroperationen und Begleiterkrankungen) auf die Testergebnisse in späteren Berechnungen zu betrachten, wurden diese absichtlich nicht konstant gehalten.
- **Innere Erkrankungen**  
Die Rekrutierung der Herz- und Krebspatienten erfolgte unmittelbar in Rehasportgruppen.
- **Orthopädie**  
Die Probanden mit Bandscheibenvorfällen und Kniearthrosen waren in Gesundheitssportgruppen aktiv. Personen mit Hüfttotalendoprothesen wurden aus einer Hüftsportgruppe in Reutlingen rekrutiert.
- **„Körperlich Gesunde“**  
Diese Personen wurden aus verschiedenen Präventivsportgruppen, u. a. der Gruppe des Instituts für Sport- und Präventivmedizin der Universität des Saarlandes, rekrutiert.

Für die Untersuchungen wurden folgende Ausschlusskriterien festgelegt:

1. Gering belastbare Herzpatienten (Thompson et al., 2013)
2. Notwendigkeit von Gehhilfen (z. B. Krücken, Rollator)
3. Akute (fiebrhafte) Erkrankungen und/oder Verletzungen

### 5.1.3 Verwendete Testverfahren

#### 5.1.3.1 Fragebögen

##### Mobilitäts-Funktionalitäts-Index (MFI)

Der MFI ist ein Fragebogen, der sich am theoretischen Modell der Mobilität im Sinne der ICF orientiert. Daher wurde dieser als Validitätskriterium des TUG<sub>plus</sub> herangezogen (Kap. 5.6.1). Der MFI beinhaltet u. a. Fragen zu Alltagsbewegungen (Fähigkeit kurze Strecken zu gehen, Gegenstände handzuhaben, leichte und schwere Gegenstände zu tragen), die auch im TUG<sub>plus</sub> integriert sind.

Der Fragebogen erfasst die „körperliche Aktivität von älteren Menschen (60 bis 85 Jahren)“ und soll Aussagen über den „Lebensstil, zur Aktivität und Inaktivität unter Berücksichtigung der individuellen Mobilität“ zulassen (Lange, 2013, S. 148). Hierbei werden 15 Items mittels fünfstufiger Likert-Skala bewertet.

Das Ergebnis bildet ein erreichter Indexwert (Maximalpunktzahl: 75). Da diese Werte jedoch noch nicht normiert wurden, lassen sich nur indirekt Rückschlüsse für das motorische Fähigkeitsniveau ableiten (Lange, 2013, S. 150). Daher wurden lediglich die erreichten Punktwerte aller MFI-Items als Summe für spätere Berechnungen verwendet.

##### Subjektiver Anstrengungsgrad

Borg (1982) entwickelte eine Skala zur Ermittlung des subjektiven Anstrengungsempfindens (Tab. 15) während fahrradergometrischen Belastungen. Die Skala reicht von sechs (sehr, sehr leicht) bis 20 (sehr, sehr schwer). Die Intention war, die Skalenstufe als Prädiktor für den Belastungspuls (Herzfrequenz: 60 - 200 Schläge pro Minute) einzusetzen (Borg, 1982, S. 379).

Die Skala diente zur Überprüfung des subjektiven Anstrengungsempfindens bei der Untersuchung zur Parallel-Test-Reliabilität (Kap. 5.4.2) und als Ausbelastungskriterium bei der Fahrradergometrie (Kap. 5.6.1).

*Tabelle 15: Subjektiver Anstrengungsgrad (Borg-Skala 6 - 20; Borg, 1982, S. 379)*

Borg-Wert	Interpretation
6	
7	sehr, sehr leicht
8	
9	sehr leicht
10	
11	ziemlich leicht
12	
13	etwas schwer
14	
15	schwer
16	
17	sehr schwer
18	
19	sehr, sehr schwer
20	

#### 5.1.3.2 Labortest

##### Fahrradergometrie

Die maximale körperliche Leistungsfähigkeit (W/kg KG) auf dem Ergometer diene als weiteres externes Validitätskriterium, da diese häufig als internationales Maß in Validitätsstudien Anwendung findet. Getestet wurde auf einem drehzahlunabhängigen Fahrradergometer (Ergo-Fit). Die Durchführung richtete sich nach den Empfehlungen von Kindermann (1987). Vor der Belastung wurde die Herzfrequenz mittels Pulsuhr registriert. Anschließend begann die Belastung bei 50 Watt und wurde jeweils nach drei Minuten um 50 Watt erhöht. Gesteigert wurde bis zur subjektiven Erschöpfung bzw. bis mindestens zur per Faustformel berechneten, altersgemäßen, maximalen Herzfrequenz (200 Schläge pro Minute minus das Lebensalter; Such, & Meyer, 2010, S. 310 f.; Kindermann, 1987, S. 9). Am Ende jeder Stufe wurden Herzfrequenz und subjektives Anstrengungsempfinden (Borg-Skala 6 bis 20) gemessen. Bei Belastungsabbruch wurde die maximal erreichte Herzfrequenz ( $HF_{peak}$ ; Such, & Meyer, 2010), das subjektive Anstrengungsempfinden und die Abbruchgründe protokolliert (Kap. 5.6.1).

### 5.1.3.3 Feldtests

#### Timed Up and Go Test plus (TUG<sub>plus</sub>)

Das weiterentwickelte Screening (s. Abb. 17, S. 62) wurde in allen Untersuchungen durchgeführt. Die weiteren Testverfahren wurden zusätzlich zur Bestimmung der Gütekriterien und diagnostischen Qualität verwendet.

#### Timed up and Go Test (TUG)

Der Test wurde zur Überprüfung der Anwendbarkeit mit verschiedenen Alters- und Personengruppen in einer Vor- bzw. Paralleluntersuchung durchgeführt (Kap. 5.2). Weiterhin galt der TUG als Kriterium zur Überprüfung der Validität (Kap. 5.6.1), da dieser die Grundlage des zu evaluierenden Verfahrens darstellt. Der Ablauf (s. Kap. 4.1, S. 47) und die Instruktionen (schnellstmöglich) waren standardisiert.

Da alle Elemente des TUG, wenn auch in abgewandelter Form (Aufstehen von der Bank statt Aufstehen vom Stuhl), im TUG<sub>plus</sub> enthalten sind und der gleiche Messparameter (Zeit) verwendet wird, ist eine hohe Korrelation beider Testverfahren anzunehmen.

Der Test diente weiterhin zur Verifizierung von Mobilitätsproblemen (Kap. 5.7.2.5.1).

#### Motorische Basisdiagnostik (MBD)

Die MBD ist ein Test zur Beurteilung der motorischen Leistungsfähigkeit (Wydra, 1996). Der Test orientiert sich am Modell der motorischen Fähigkeiten von Bös und Mechling (1983) und beinhaltet dichotome Items zur Ausdauer (stufenförmige Belastung mit den Geschwindigkeiten 5 bis 7 km/h, optische Geschwindigkeitsvorgabe durch eine Lauflichtsteuerung), Kraft (Schulterwegdrücken, Hüftstrecken, Einbeinaufstehen, Rumpfaufrichten), Beweglichkeit (Rumpfflexion und -extension, Bein- strecken, Anfersen, Ausschultern) und Koordination (Achterkreisen, Balancieren und Ballprellen, Ball umgreifen, Arm-Bein-Koordination an der Wand entlang), die später als Summenscores (erreichte Punktwerte in den vier motorischen Fähigkeitsbereichen) und Gesamtsummenscore (erreichter Punktwert über alle Items) die Messkriterien stellten.

Die einzelnen Items wurden von jedem Probanden nacheinander durchgeführt.

Die MBD wurde zur Überprüfung der Kriteriumsvalidität durchgeführt (Kap. 5.6.1). Der TUG<sub>plus</sub> besteht aus Testaufgaben, die jeweils von mehreren motorischen Fähigkeiten limitiert werden. Um diese Fähigkei-



ten mit den fertigungsorientierten Testaufgaben des Screenings in Verbindung zu bringen, wurde die MBD verwendet. Zudem wurde diese zur Verifizierung motorischer Probleme in der diagnostischen Untersuchung (Kap. 5.7.2.5.1) durchgeführt.

### Handkraft

Die Messung der Handkraft dient häufig als ein Indikator für den Gesundheitszustand von Personen (Giampaoli et al., 1999; Soety, Faruquee, Lahiry, Chaklader, & Yasmin, 2014). Zudem gilt diese auch als Prädiktor von Funktionseinschränkungen (Rantanen, Guralnik, Foley, Masaki, Leveille, Curb, & White, 1999, S. 558).

Die Messung der isometrischen Handkraft der linken und rechten Hand erfolgte mittels hydraulischen Handdynamometers (Jamar Hand Dynamometer). Dieses Gerät wird international zur Beurteilung der Handkraft am häufigsten verwendet (Roberts, Denison, Martin, Patel, Syddall, Cooper, & Sayer, 2011, S. 424).

Der Griff des Dynamometers lässt sich in fünf Positionen einstellen und wurde zunächst, an die Handgröße angepasst, voreingestellt. Bei ausgestreckter Hand sollten sich alle Fingerglieder über dem Griffbereich befinden (Abb. 19a). Dies entsprach bei allen Probanden der zweiten Griffposition. Diese Position wird auch in Studien aufgrund der höchsten Reliabilität am häufigsten verwendet (Roberts et al., 2011, S. 425). Die größte Kraftentfaltung ist im Stehen bei 90° Armflexion im Ellbogengelenk möglich (Balogun, Akomolafe, & Amusa, 1991; Mathiowetz, Rennells, & Donahoe, 1985). Diese Position wurde deshalb mit zusätzlich adduziertem Ellbogen als Ausgangsposition gewählt (Abb. 19b).

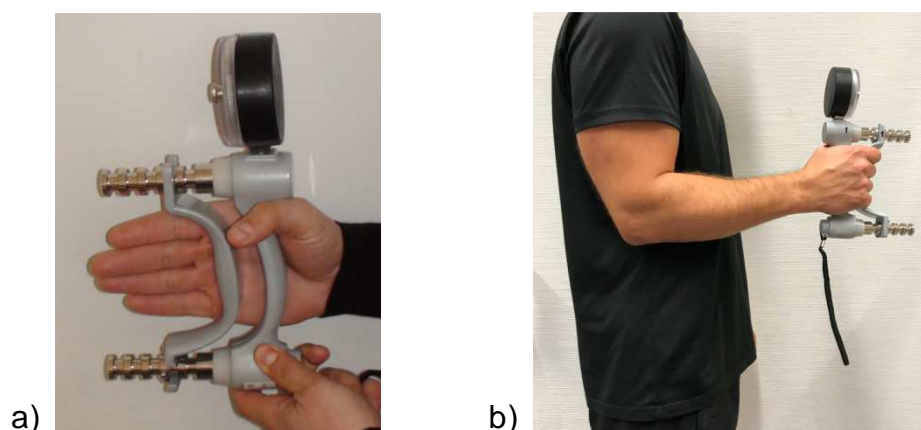


Abbildung 19: Messung der Handkraft: a) Handhaltung, b) Armhaltung

Die Probanden hatten die Aufgabe, unter Motivation durch den Testleiter, die Hand maximal fest zusammen zu drücken. Anschließend wurde der erreichte Wert in Kilogramm protokolliert. Jede Hand wurde abwechselnd zweimal getestet (Testreihenfolge: rechts, links, rechts, links). Anschließend wurden die Mittelwerte beider Seiten für weitere Berechnungen verwendet.

Da der TUG<sub>plus</sub> neben koordinativen Aufgaben, auch Items, die ein gewisses Maß an Kraft voraussetzen (z. B. Kasten umstoßen, Kasten tragen) beinhaltet, diente die Handkraft als weiteres Validitätskriterium (Kap. 5.6.1).

#### 5.1.4 Statistik

Die in den Untersuchungen gewonnenen Daten wurden mit dem Programm IBM SPSS, Version Statistics 22 und faktorenanalytische Betrachtungen mittels Amos 25 ausgewertet.

Angaben zu Häufigkeitsverteilungen (n; %), Mittelwerten (M) und Standardabweichungen (SD) dienten als deskriptive Maße.

Die Überprüfung von Normalverteilungen ( $p > 0,05$ ) der abhängigen Variablen und der Residuen erfolgte mittels Histogramme und dem Kolmogorov-Smirnov-Test.

Unterschiedshypothesen von zwei unabhängigen Gruppen wurden bei Normalverteilung mittels t-Test geprüft. Bei vorliegender Nicht-Normalverteilung wurde der Mann-Whitney-U-Test verwendet. Zur Berechnung von Unterschieden bei mehr als zwei Gruppen wurde bei Datensätzen mit normalverteilten Residuen die einfaktorielle Varianzanalyse verwendet. Hierbei erfolgte zunächst eine Überprüfung der Homogenität der Varianzen mittels Levene-Test. Bei Varianzhomogenität ( $p > 0,05$ ) wurde der Scheffé- und bei Varianzheterogenität ( $p \leq 0,05$ ) der Tamhane-T2-Test verwendet.

Zusammenhangshypothesen wurden bei normalverteilten, abhängigen Variablen mit dem Produkt-Moment-Korrelationskoeffizient nach Pearson und bei nicht normalverteilten Daten mit dem Spearman-Rangkorrelationskoeffizienten berechnet.

Hinsichtlich der Interpretation der Korrelationskoeffizienten wurde die Abstufung von Bös, Hänsel und Schott (2004, S. 169) verwendet (Tab. 16).

*Tabelle 16: Korrelationskoeffizient ( $r$ ) und entsprechende Interpretation (Bös, Hänsel, & Schott, 2004, S. 169)*

Korrelationskoeffizient ( $r$ )	Interpretation
0,00	Kein Zusammenhang
$0,01 \leq 0,39$	Niedriger Zusammenhang
$0,40 \leq 0,69$	Mittlerer Zusammenhang
$0,70 \leq 0,99$	Hoher Zusammenhang
1,00	Perfekter Zusammenhang

Das Signifikanzniveau wurde folgendermaßen abgestuft (Tab. 17).

*Tabelle 17: Signifikanzniveau ( $p$ ) und entsprechende Interpretation (Clauß, & Ebner, 1975, S. 189)*

Signifikanz ( $p$ )	Interpretation
$> 0,05$	Nicht signifikant (n.s.)
$\leq 0,05$	Signifikant (*)
$\leq 0,01$	Sehr signifikant (**)
$\leq 0,001$	Hoch signifikant (***)

Zur Überprüfung der Effektstärke bei signifikanten Ergebnissen dienen Eta-Quadrat ( $\eta^2$ ) und Cohens  $d$  (Tab. 18).

*Tabelle 18: Effektstärkemaße Eta-Quadrat ( $\eta^2$ ) und Cohens  $d$  (Cohen, 1988)*

Eta-Quadrat ( $\eta^2$ )	Cohens $d$	Effekt
$0,01 < 0,06$	$< 0,5$	Kleiner Effekt
$0,06 < 0,14$	$0,5 \leq 0,8$	Mittlerer Effekt
$\geq 0,14$	$> 0,8$	Großer Effekt

Zur Berechnung faktorenanalytischer Fragestellungen (Kap. 5.6.2) wurden zunächst Durchführungsvoraussetzungen überprüft: Die Kommunalitäten ( $h_i^2$ ) der Items (Varianzaufklärung durch Faktorenextraktion) dienten zur Bestimmung der Itemqualitäten (Bühner, 2004, S. 168). Die Itemkorrelationen wurden mittels Kaiser-Meyer-Olkin-Koeffizient (KMO: Indikator zur Eignung der Items) und Bartlett-Test (Test auf Nullkorrelation der Korrelationen) getestet und auf Multikollinearität (starke Korrelation der Variablen) überprüft (Bühner, 2004, S. 170 f.; Pospeschill, 2010, S. 208).

Anschließend erfolgten explorative (EFA: Hauptkomponentenanalyse zur Reproduktion ursprünglicher Itemkorrelationen) und konfirmatorische Faktorenanalysen (KFA: Maximum-Likelihood-Schätzung zur Faktorenschätzung größtmöglicher Varianzaufklärung; Bühner, 2004, S. 158, 160). Mittels Chiquadrat-Test ( $\chi^2$ ) und ergänzenden Fit Indizes

sollte im Weiteren die Anpassung des in der KFA aufgestellten Modells überprüft werden (Tab. 19). Anhand des  $\chi^2$ -Wertes bedeuten nicht signifikante Werte eine gute Anpassung des Modells (Pospeschill, 2010, S. 221). Die Fit Indizes Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI) und Tucker Lewis Index (TLI) setzen das getestete Modell mit einem unabhängigen Modell in Zusammenhang und dienen zur weiteren spezifischeren Überprüfung der Modellanpassung (Tab. 19; Bühner, 2004, S. 204; Pospeschill, 2010, S. 220 f.).

*Tabelle 19: Fit Indizes: Formeln und Grenzwerte für die Qualität von Modellanpassungen (Hu, & Bentler, 1999, S. 5 ff.; Bühner, 2004, S. 205; Pospeschill, 2010, S. 220 f.)*

<b>Fit Index</b>	<b>Gute Modellanpassung</b>
Root Mean Square Error of Approximation (RMSEA)	$\leq 0,06$ ( $n > 250$ ) $\leq 0,08$ ( $n < 250$ )
Comparative Fit Index (CFI)	$\geq 0,95$
Tucker Lewis Index (TLI)	$\geq 0,95$

Zur Berechnung von Grenzwerten wurden Ergebnisverteilungen mittels eines Boxplots dargestellt (Kap. 5.7.1). ROC-Analysen und die Qualitätskriterien Sensitivität, Spezifität, prädiktive Werte, Trennschärfe, Trefferquote und Wahrscheinlichkeitsverhältnisse dienten zur Bestimmung der diagnostischen Güte (Kap. 5.7.2).

Die bei den jeweiligen Teiluntersuchungen verwendeten spezifischen Verfahren werden in den einzelnen Kapiteln aufgeführt.

## 5.2 Vor- bzw. Paralleluntersuchung – Klassischer TUG

Der TUG besitzt aufgrund der Verwendung nur eines Messparameters – Zeit bzw. Geschwindigkeit – eine hohe Testökonomie. Dadurch kommt der Testinstruktion eine besondere Bedeutung zu. Wenn die Instruktion nicht eindeutig auf die maximal mögliche Geschwindigkeit abzielt, resultiert hieraus eine erhebliche Störgröße (s. Tab. 9, S. 51). Dies bestätigten auch einige im Rahmen dieser Studie durchgeführten Vorversuche, bei denen der Test mit den Instruktionen „schnellstmöglich“ und „schnellstmöglich und sicher“ durchgeführt wurde.

Die Probanden (62 bis 84 Jahre) bei der Studie von Pertille et al. (2012) wurden trotz ihres Alters mit der Anweisung „schnellstmöglich“ instruiert und erreichten damit eine Zeit zwischen 6 und 7 s. Obwohl die Probanden bei Kumar et al. (2015) deutlich jünger waren (38 bis 55 Jahre), erreichten sie mit der Anweisung „schnellstmöglich und sicher“ lediglich einen vergleichbaren Geschwindigkeitsbereich (s. Tab. 9, S. 51).

Also sollte vor allem bei zeitbezogenen Tests auf die Instruktion besonderen Wert gelegt werden.

Im Folgenden sollten die Testzeiten des TUG bei einer möglichst großen Stichprobe mit der Instruktion „schnellstmöglich“ und die in der Literatur verwendeten Grenzwerte überprüft werden.

### 5.2.1 Fragestellungen

Im Folgenden sollen Testzeiten in Abhängigkeit vom Alter und Gesundheitsstatus (Patienten und „Gesunde“) gemessen werden.

Hierzu stehen zwei Fragestellungen im Fokus:

1. Liefert der TUG auch bei jüngeren Personengruppen (< 60 Jahre) plausible Ergebnisse?
2. Unterscheiden sich die Testzeiten bei verschiedenen Krankheitsbildern und „gesunden“ Personen?
3. Zeigen die gemessenen Testzeiten geschlechts- und altersabhängige Unterschiede?

### 5.2.2 Methodik – Personenstichprobe und Untersuchungsablauf

An den Testungen nahmen 1023 Personen (♀ 54,4 %; Alter  $56 \pm 11$  Jahre) teil (Tab. 20).

*Tabelle 20: TUG: Anthropometrische Daten der getesteten Personengruppen beim TUG ( $M \pm SD$ ;  $n = 1023$ )*

Personengruppe (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
<b>HNO</b> (310)	$52 \pm 9$	$172 \pm 9$	$78 \pm 17$
<b>Neurologie</b> (222)	$54 \pm 10$	$173 \pm 8$	$80 \pm 18$
<b>Innere</b> (227)	$63 \pm 10$	$172 \pm 29$	$74 \pm 16$
<b>Orthopädie</b> (127)	$57 \pm 13$	$170 \pm 10$	$76 \pm 14$
<b>„Gesunde“</b> (137)	$55 \pm 11$	$174 \pm 10$	$76 \pm 14$

Der TUG wurde standardisiert mit der Instruktion „schnellstmöglich“ von den Probanden durchgeführt (s. Abb. 14, S. 47).

### 5.2.3 Hypothesen

- Hypothese 1: Die Testzeiten (s) unterscheiden sich in Abhängigkeit von den Personengruppen ( $p \leq 0,05$ ;  $\eta^2 \geq 0,06$ ;  $d \geq 0,5$ ).
- Hypothese 2: Die Testzeiten (s) unterscheiden sich in Abhängigkeit vom Geschlecht und Alter der Probanden ( $p \leq 0,05$ ;  $\eta^2 \geq 0,06$ ;  $d \geq 0,5$ ).

### 5.2.4 Statistik

Die Berechnung von Unterschieden der mittleren Zeiten bei den Personengruppen (normalverteilte Residuen) erfolgte mittels einfaktorieller Varianzanalyse mit einem Tamhane-T2-Test (Varianzheterogenität) als post-hoc Verfahren.

Die Überprüfung geschlechtsabhängiger Unterschiede erfolgte mittels t-Test. Hinsichtlich altersabhängiger Unterschiede wurde zur besseren Übersicht bei jeder Personen- bzw. Patientengruppe eine Einteilung des Alters in Quartile vorgenommen. Die Zeiten in Abhängigkeit der Altersquartile (normalverteilte Residuen) wurden anschließend mittels einfaktorieller Varianzanalyse und Scheffé-Test (Varianzhomogenität) auf Unterschiede geprüft.

Zur Beurteilung der Varianzaufklärung wurde das Effektstärkemaß Eta-Quadrat ( $\eta^2$ ) und für die Unterschiede zwischen jeweils zwei Mittelwerten Cohens d verwendet.

### 5.2.5 Ergebnisse

#### *Personengruppen*

Tabelle 21 zeigt die durchschnittlich benötigten Testzeiten. Durchschnittlich durchliefen die Probanden den Parcours in  $6,5 \pm 1,5$  s.

*Tabelle 21: Testzeiten (s) des TUG bei verschiedenen Personengruppen ( $M \pm SD$ ;  $n = 1023$ )*

	Personengruppe					
	HNO (n = 310)	Neurologie (n = 222)	Innere (n = 227)	Orthopädie (n = 127)	„Gesunde“ (n = 137)	Gesamt (n = 1023)
<b>Zeit</b> (s)	$6,2 \pm 1,5$	$6,9 \pm 1,5$	$6,4 \pm 1,4$	$6,6 \pm 1,7$	$5,5 \pm 1,2$	$6,5 \pm 1,5$

Die mittleren Testzeiten unterschieden sich zwischen den Gruppen ( $\eta^2 = 0,08$ ;  $F = 21,3$ ).

„Körperlich Gesunde“ durchliefen den TUG erwartungsgemäß schneller ( $p \leq 0,001$ ) als Patienten der HNO ( $d = 0,52$ ), Neurologie ( $d = 1,0$ ), inneren Medizin ( $d = 0,69$ ) und Orthopädie ( $d = 0,75$ ).

Neurologische Patienten benötigten im Mittel die meiste Zeit. Sie waren langsamer ( $p \leq 0,001$ ) als Patienten mit inneren ( $d = 0,35$ ) und HNO-Erkrankungen ( $d = 0,47$ ; Abb. 20).

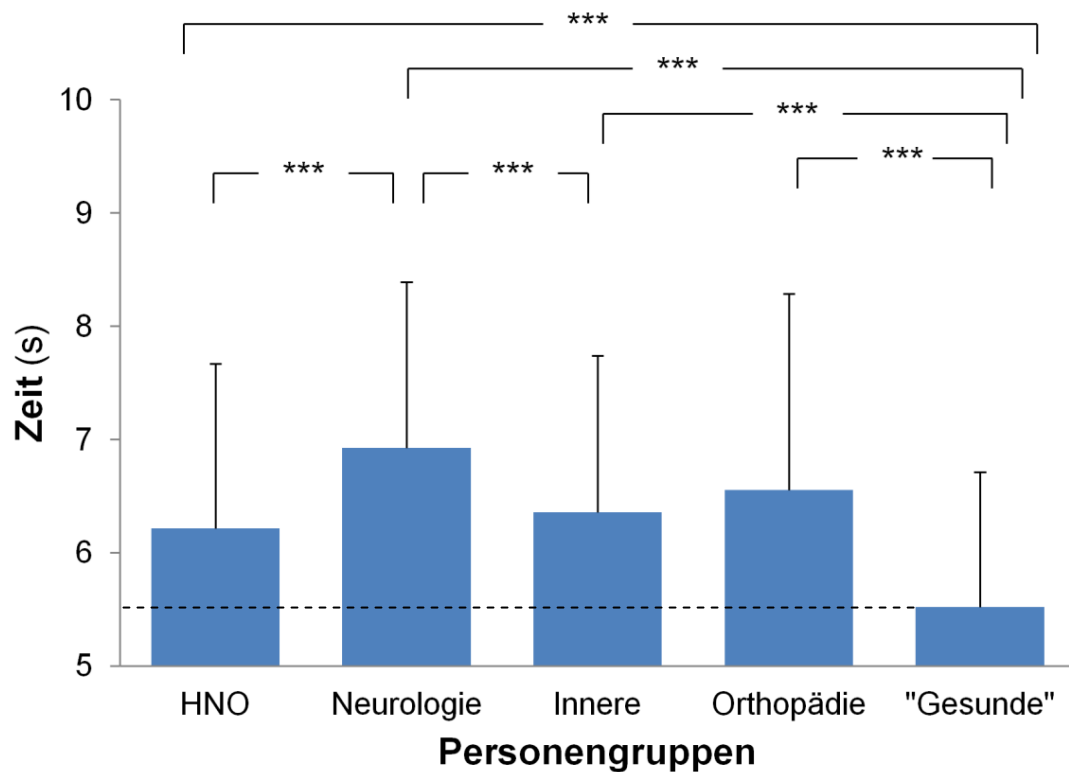


Abbildung 20: Patientengruppenabhängige Testzeiten (s) beim TUG in Bezug zu den „Körperlich Gesunden“ ( $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ ; gestrichelte Linie = 5,5 s;  $n = 1023$ )

#### Geschlecht und Alter

Es ergaben sich keine geschlechtsabhängigen Testzeiten ( $p > 0,05$ ; ♀  $6,3 \pm 1,6$  s; ♂  $6,4 \pm 1,4$  s).

Die jüngste Person (HNO-Patientin, 18 Jahre) benötigte beim TUG 5,0 s. Die älteste Person (Orthopädie-Patientin, 87 Jahre) absolvierte den Test in 9,8 s.

Anhand der Altersverteilung ergaben sich bei den Personengruppen unterschiedliche Quartile und dementsprechend verschiedene mittlere Testzeiten (Tab. 22).



*Tabelle 22: Altersspezifische Testzeiten (s; n = 1023)*

Stichprobe (n)	Altersquartile (Jahre)	Zeit (s)
Gesamtstichprobe (1023)	≤ 49	6,2 ± 1,4
	50 - 55	6,2 ± 1,5
	56 - 61	6,2 ± 1,5
	> 61	6,7 ± 1,5
	gesamt	6,3 ± 1,5
Neurologie (222)	≤ 48	6,7 ± 1,4
	49 - 54	6,8 ± 1,6
	55 - 60	6,8 ± 1,3
	> 60	7,3 ± 1,5
	gesamt	6,9 ± 1,5
HNO (310)	≤ 47	5,8 ± 1,2
	48 - 54	6,1 ± 1,2
	55 - 58	6,3 ± 1,6
	> 58	6,9 ± 1,7
	gesamt	6,2 ± 1,5
Innere (227)	≤ 56	5,8 ± 1,4
	57 - 63	6,0 ± 1,1
	64 - 70	6,7 ± 1,1
	> 70	7,0 ± 1,5
	gesamt	6,4 ± 1,4
Orthopädie (127)	≤ 49	7,1 ± 1,8
	50 - 57	6,4 ± 2,1
	58 - 65	5,9 ± 1,5
	> 65	6,7 ± 1,4
	gesamt	6,6 ± 1,7
„Gesunde“ (137)	≤ 49	5,5 ± 0,8
	50 - 55	5,7 ± 1,4
	56 - 62	5,5 ± 1,4
	> 63	5,4 ± 1,2
	gesamt	5,5 ± 1,2

Die Altersquartile der Gesamtstichprobe erklären 2 % der Gesamtvarianz ( $\eta^2 = 0,02$ ). Personen mit einem Alter über 62 Jahre durchliefen den TUG langsamer ( $6,7 \pm 1,5$  s) als die jüngeren Personengruppen, die sich bzgl. der Testzeit nicht weiter voneinander unterschieden ( $\leq 49$  Jahre:  $6,2 \pm 1,4$  s;  $p \leq 0,001$ ;  $d = 0,35$ ; 50 bis 55 Jahre:  $6,2 \pm 1,5$  s;  $p \leq 0,01$ ;  $d = 0,33$ ; 56 bis 61 Jahre:  $6,2 \pm 1,4$  s;  $p \leq 0,01$ ;  $d = 0,35$ ;  $F = 7,4$ ; Abb. 21).

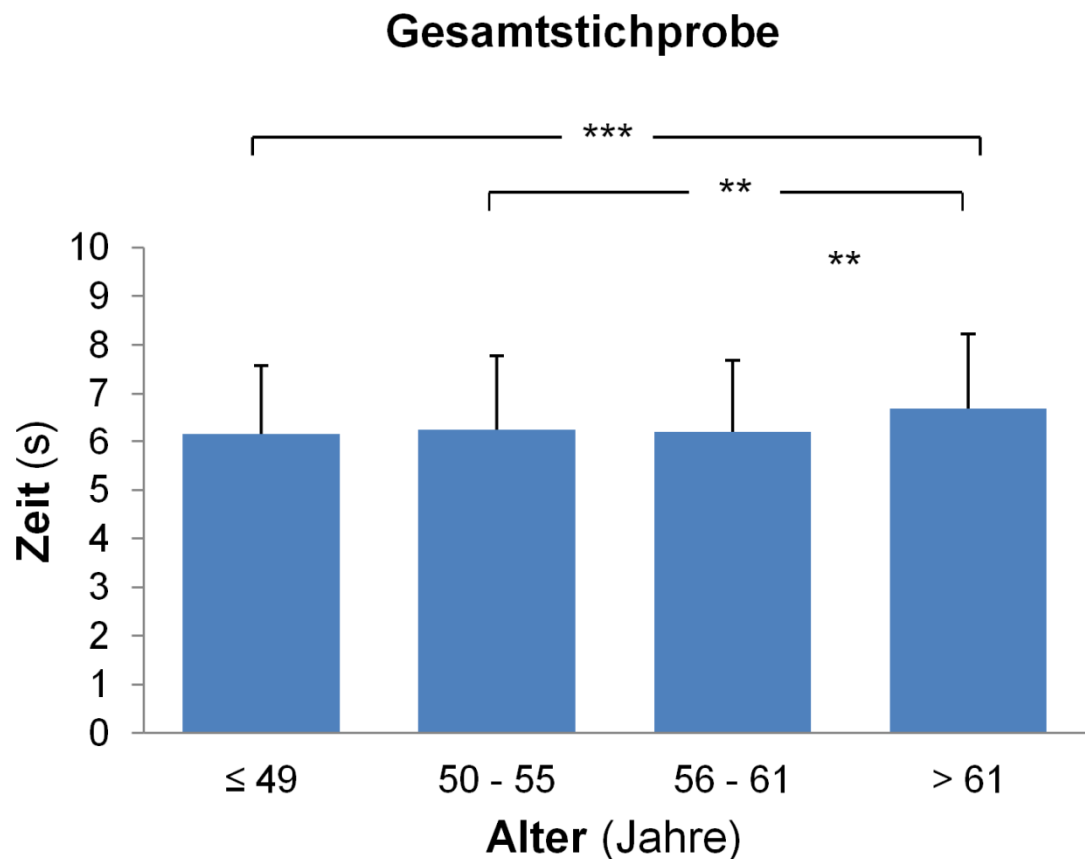


Abbildung 21: Unterschiede der Altersquartile hinsichtlich der Testzeiten (s) beim TUG ( $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ ; \*\* =  $p \leq 0,01$ ;  $n = 1023$ )

„Gesunde“ Probanden benötigten im Mittel weniger als 6 s für den TUG ( $\leq 49$  Jahre:  $5,5 \pm 0,8$  s; 50 bis 55 Jahre:  $5,7 \pm 1,4$  s; 56 bis 62 Jahre:  $5,5 \pm 1,4$  s;  $\geq 63$  Jahre:  $5,4 \pm 1,2$  s). Es ergaben sich hinsichtlich des Alters keine unterschiedlichen Testzeiten ( $p > 0,05$ ; Abb. 22).

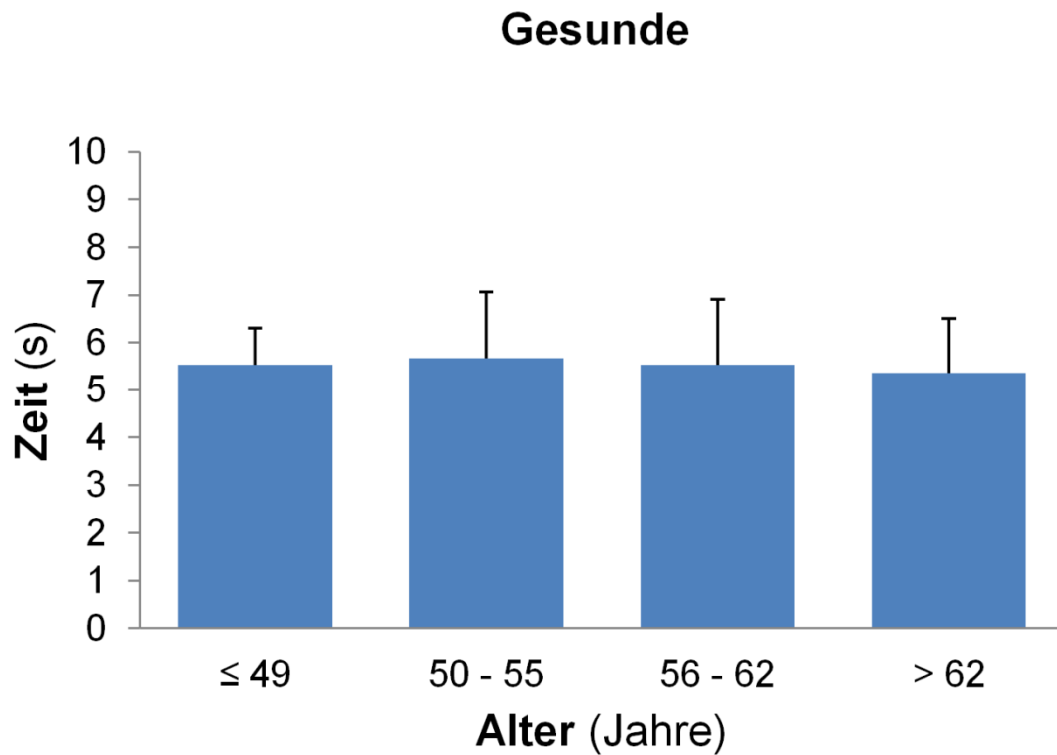


Abbildung 22: Altersabhängige Unterschiede der Testzeiten (s) beim TUG bei „Körperlich Gesunden“ ( $M \pm SD$ ;  $n = 137$ )

Hinsichtlich der Testzeiten bestehen signifikante Unterschiede zwischen den Altersspannen ( $\eta^2 = 0,07$ ;  $F = 7,9$ ). Personen jünger als 47 Jahre benötigten die geringste Zeit ( $5,8 \pm 1,2$  s). Diese Probanden und Patienten im Alter zwischen 48 und 54 Jahren waren schneller als Personen in einem Alter von 59 Jahren oder älter ( $6,9 \pm 1,7$  s;  $p \leq 0,001$ ;  $d = 0,70$  bzw.  $p \leq 0,01$ ;  $d = 0,54$ ; Abb. 23).

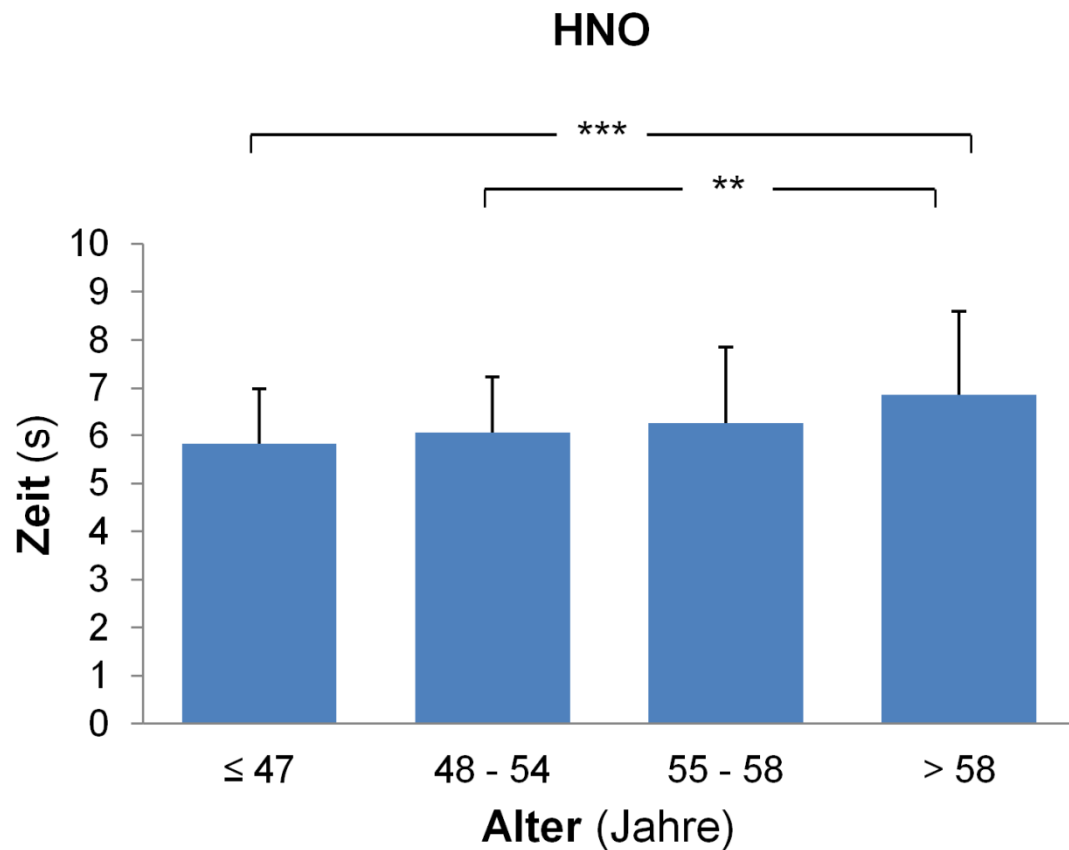


Abbildung 23: Altersabhängige Unterschiede der Testzeiten (s) beim TUG bei HNO-Patienten ( $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ ; \*\* =  $p \leq 0,01$ ;  $n = 310$ )

Neurologische Patienten benötigten im Mittel zwischen 6,7 und 7,3 s ( $\leq 48$  Jahre:  $6,7 \pm 1,4$  s; 49 bis 54 Jahre:  $6,8 \pm 1,6$  s; 55 bis 60 Jahre:  $6,8 \pm 1,3$  s;  $\geq 61$  Jahre:  $7,3 \pm 1,5$  s. Hier zeigen sich jedoch keine Unterschiede ( $p > 0,05$ ;  $\eta^2 = 0,03$ ; Abb. 24).

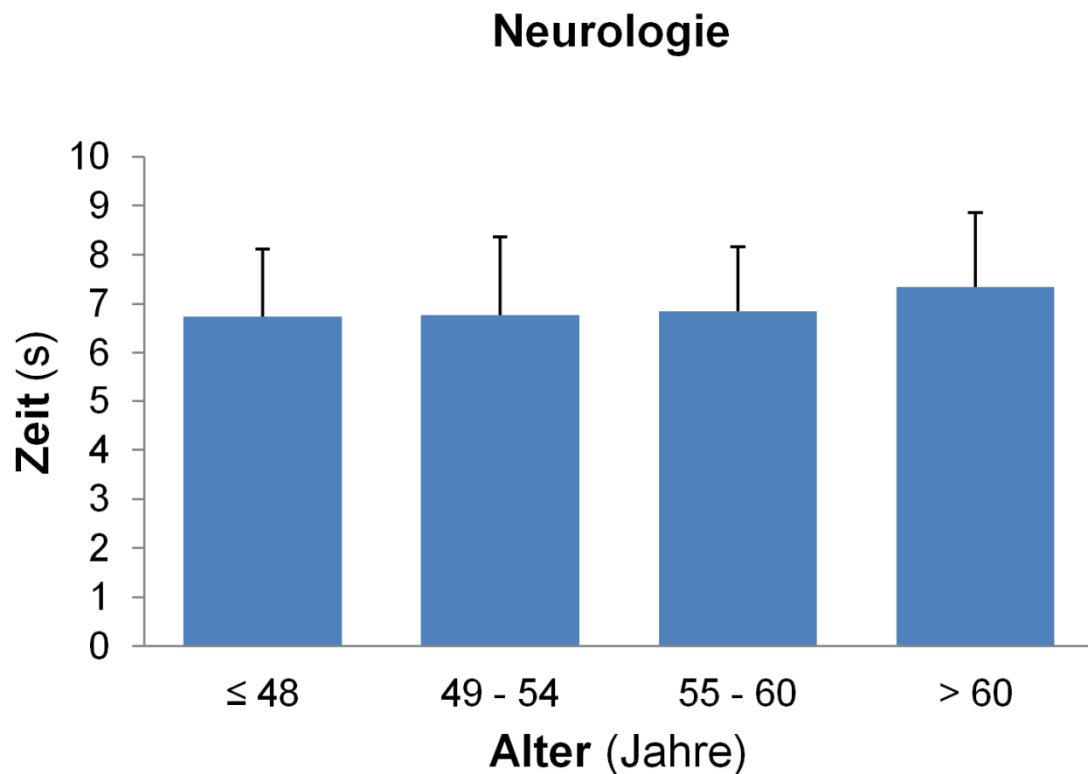


Abbildung 24: Altersabhängige Unterschiede der Testzeiten (s) beim TUG bei neurologischen Patienten ( $M \pm SD$ ;  $n = 222$ )

Bei Patienten mit inneren Erkrankungen zeigt das Alter einen höheren Einfluss auf die Testzeit ( $\eta^2 = 0,13$ ). Zwischen den Altersspannen ergaben sich signifikante Unterschiede ( $F = 10,6$ ). Personen jünger als 56 Jahre durchliefen den TUG schneller ( $5,8 \pm 1,4$  s) als Personen im Alter zwischen 64 und 70 Jahren ( $6,7 \pm 1,1$  s;  $p \leq 0,01$ ;  $d = 0,72$ ) und älter als 71 Jahren ( $7,0 \pm 1,5$  s;  $p \leq 0,001$ ;  $d = 0,83$ ). Weiterhin waren Patienten im Alter zwischen 57 und 63 Jahren ( $6,0 \pm 1,1$  s) schneller als Probanden zwischen 64 und 70 Jahren ( $p \leq 0,05$ ;  $d = 0,64$ ) und älter als 71 Jahren ( $p \leq 0,01$ ;  $d = 0,76$ ; Abb. 25).

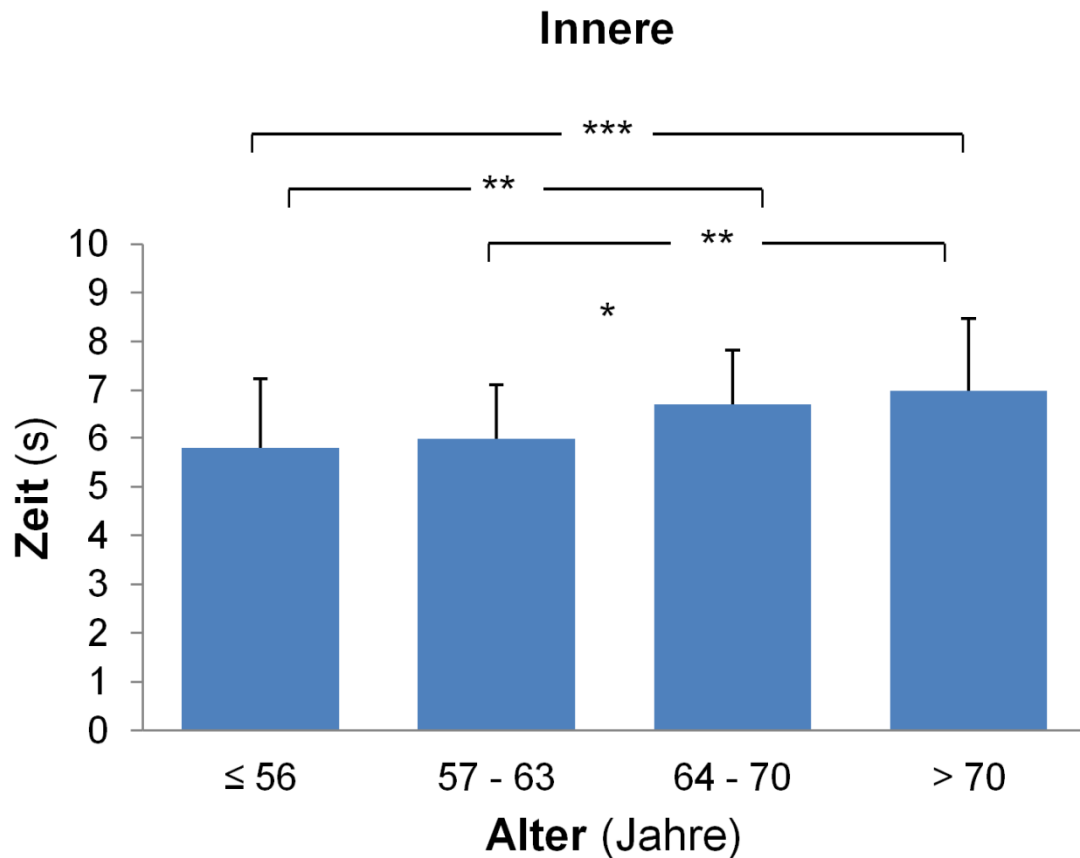


Abbildung 25: Altersabhängige Unterschiede der Testzeiten (s) des TUG bei Personen mit inneren Erkrankungen ( $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ ; \*\* =  $p \leq 0,01$ ; \* =  $p \leq 0,05$ ;  $n = 227$ )

Bei den orthopädischen Patienten waren, entgegen der Erwartung, die jüngsten Probanden am langsamsten ( $7,1 \pm 1,8$  s). Bis zu einem Alter von 65 Jahren nahmen die Geschwindigkeiten zu (50 bis 57 Jahre:  $6,4 \pm 2,1$  s; 58 bis 65 Jahre:  $5,9 \pm 1,5$  s). Probanden über 65 Jahre benötigten wieder länger zur Absolvierung der Items ( $6,7 \pm 1,4$  s). Die Zeitunterschiede zwischen den Altersspannen waren jedoch nicht signifikant (Abb. 26).

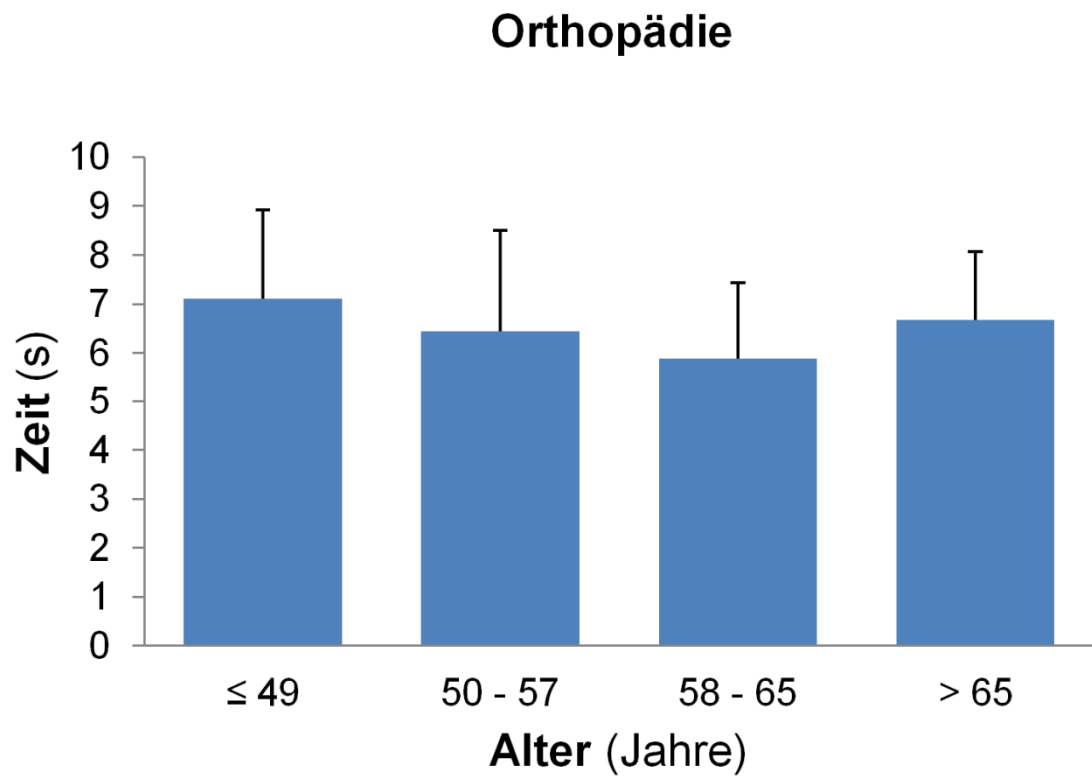


Abbildung 26: Altersabhängige Unterschiede der Testzeiten (s) beim TUG bei orthopädischen Patienten ( $M \pm SD$ ;  $n = 127$ )

## 5.3 Evaluation des TUG<sub>plus</sub> anhand verschiedener Alters- und Personengruppen

### 5.3.1 Fragestellungen

Der klassische TUG wurde in der Vergangenheit auch schon mit verschiedenen Alters- und Personengruppen durchgeführt, bei denen sich unterschiedliche Normwertaufstellungen finden (s. Tab. 8, S. 48).

Deshalb sollte hier überprüft werden, ob anhand des TUG<sub>plus</sub> ebenfalls adressatenspezifische Testzeiten ermittelt werden können und die Schwierigkeit der Items bei verschiedenen Personengruppen ausreichend differenziert. Da eine geringe Testzeit eine ökonomische Durchführung der Items erfordert, sollte zudem die Frage beantwortet werden, ob ein Zusammenhang zwischen der Beurteilung der Ausführungsökonomie (s. Tab. 14, S. 66) und der Testzeit besteht.

### 5.3.2 Methodik

#### 5.3.2.1 Personenstichprobe

An der Evaluation nahmen insgesamt 764 Probanden unterschiedlicher Patienten- bzw. Personengruppen teil (Tab. 23).

*Tabelle 23: Anthropometrische Daten der Gesamtstichprobe beim TUG<sub>plus</sub>, differenziert nach Personengruppen ( $M \pm SD$ ;  $n = 764$ )*

Personengruppen (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Neurologie (169)	♀ (71)	51 ± 12	167 ± 6	72 ± 15
	♂ (98)	55 ± 10	176 ± 12	87 ± 17
HNO (224)	♀ (118)	52 ± 10	166 ± 6	71 ± 14
	♂ (106)	54 ± 10	180 ± 7	90 ± 15
Innere (128)	♀ (67)	64 ± 8	163 ± 6	84 ± 14
	♂ (61)	66 ± 11	175 ± 7	66 ± 11
Orthopädie (79)	♀ (48)	61 ± 14	165 ± 7	70 ± 13
	♂ (31)	58 ± 13	177 ± 6	85 ± 14
„Gesunde“ (153)	♀ (97)	59 ± 11	165 ± 7	65 ± 14
	♂ (56)	58 ± 12	180 ± 6	83 ± 14
Sportstudierende (11)	♀ (4)	22 ± 3	167 ± 7	58 ± 8
	♂ (7)	23 ± 3	179 ± 6	77 ± 11



### 5.3.2.2 Variablenstichprobe

Der Parcours wurde von allen Probanden durchgeführt (s. Kap. 4.2.2, S. 60). Die Bewegungsausführung wurde anhand des Bewertungsschemas (s. Tab. 14, S. 66) bewertet.

### 5.3.2.3 Ablauf der Untersuchung

Nach Beantwortung der Fragen eines Anamnesebogens demonstrierte der Testleiter den Parcours und erklärte die Items. Dann absolvierte jeder Proband zunächst einen Probedurchlauf, bei dem die Items geübt wurden. Anschließend startete der Testparcours, der ohne weitere Animation schnellstmöglich durchlaufen werden sollte.

Der Testleiter notierte bei jedem Item die erfolgreiche Bewältigung und Bewegungsqualität (s. Tab. 14, S. 66;  $n = 764$ ). Nur bei Absolvierung aller Items wurde die benötigte Zeit protokolliert und für weitere Berechnungen verwendet ( $n = 670$ ).

### 5.3.3 Hypothesen

Da schon die Ergebnisse des TUG bei der Voruntersuchung und in der Literatur bei verschiedenen Personengruppen (Bischoff et al., 2003; Chiu et al., 2003; Nordin et al., 2008; Shumway-Cook et al., 2000; Lin et al., 2004; Whitney et al., 2004; Alexandre et al., 2012; Rose et al., 2002; Shimada et al., 2009; Haines et al., 2008; Kristensen et al., 2007; Savva et al., 2012), Altersgruppen (Bohannon, 2006) und hinsichtlich des Geschlechts (Thrane, Joakimsen, & Thornquist, 2007) Unterschiede zeigen, wurden hinsichtlich des TUG<sub>plus</sub> folgende Hypothesen aufgestellt.

- Hypothese 3: Männer durchlaufen den Parcours schneller als Frauen ( $p \leq 0,05$ ;  $d > 0,5$ ).
- Hypothese 4: Die Personengruppen („Gesunde“ und die jeweiligen Krankheitsbilder) unterscheiden sich hinsichtlich der benötigten Testzeit ( $p \leq 0,05$ ;  $\eta^2 > 0,06$ ;  $d > 0,5$ ).
- Hypothese 5: Die Testzeiten unterscheiden sich in Abhängigkeit vom Alter ( $p \leq 0,05$ ;  $\eta^2 > 0,06$ ;  $d > 0,5$ ).
- Hypothese 6: Zwischen den Testzeiten (s) und der ökonomischen Ausführung der Testitems (Beurteilungsbogen, s. Tab. 14, S. 66) besteht ein Zusammenhang ( $p \leq 0,05$ ;  $r > -0,40$ ).

### 5.3.4 Statistik

Zunächst wurden die Lösungshäufigkeiten und durchschnittlich benötigten Zeiten zur Absolvierung des Parcours überprüft. Hierzu wurde u. a. das Alter der Teilnehmer (Gesamtstichprobe) in Quartile eingeteilt ( $\leq 49$  Jahre, 49 bis 57 Jahre, 57 bis 63 Jahre,  $> 63$  Jahre). Anschließend erfolgte eine Betrachtung der prozentualen Lösungshäufigkeiten jedes Items.

Unterschiede zwischen den Geschlechtern bei den Gruppen Neurologie, HNO, Innere, Orthopädie und Gesunde wurden aufgrund normalverteilter Testzeiten mittels t-Tests berechnet. Bei den Sportstudierenden diente der Mann-Whitney-U-test aufgrund der geringen Stichprobengröße als Verfahren.

Zur Überprüfung der Unterschiede zwischen den Personengruppen wurde aufgrund normalverteilter Residuen der Zeiten eine einfaktorielle Varianzanalyse durchgeführt. Die Varianzen waren heterogen, so dass der Tamhane-T2-Test als post-hoc Verfahren verwendet werden konnte.

Zur Überprüfung altersabhängiger Unterschiede der Testzeiten dienten die Altersquartile der Gesamt- und Teilstichproben. Die Gruppe der Sportstudierenden wurde aufgrund hoher Altershomogenität nicht betrachtet. Aufgrund normalverteilter Residuen diente bei der Gesamtstichprobe und bei Neurologie-Patienten die einfaktorielle Varianzanalyse mit einem Tamhane-T2 als post-hoc Verfahren zur Überprüfung altersabhängiger Unterschiede. Bei den weiteren Patientengruppen sowie bei Gesunden kamen jeweils einfaktorielle Varianzanalysen mit einem Scheffé-Test zum Einsatz.

Der Zusammenhang zwischen der Testzeit und der ökonomischen Ausführung der Testitems wurde aufgrund nicht normalverteilter Daten mittels Spearman-Rangkorrelationskoeffizienten berechnet.

### 5.3.5 Ergebnisse

#### 5.3.5.1 Lösungshäufigkeiten

Von 764 Probanden lösten 88 % alle Items. Die Gruppe der Sportstudenten hatte die höchste (100 %) und HNO-Patienten mit durchschnittlich 82 % die geringste Lösungswahrscheinlichkeit. Insgesamt zeigten Frauen die meisten Probleme bei der Bewältigung des TUG<sub>plus</sub> (Tab. 24).

*Tabelle 24: Lösungshäufigkeiten der Items, differenziert nach Personengruppen (%; n = 764)*

Personengruppen (n)	Geschlecht	Alle Items erfüllt (%)
<b>Neurologie</b> (169)	♀	83
	♂	95
<b>HNO</b> (224)	♀	71
	♂	92
<b>Innere</b> (128)	♀	87
	♂	84
<b>Orthopädie</b> (79)	♀	81
	♂	97
<b>„Gesunde“</b> (153)	♀	95
	♂	100
<b>Sportstudierende</b> (11)	♀	100
	♂	100

Probanden, die ein bis drei Items des TUG<sub>plus</sub> nicht absolvieren konnten, wiesen die folgenden Beschwerden bzw. Krankheitsbilder auf:

- HNO-Patienten, die den Parcours nicht komplett durchführen konnten (♀ 29 %, ♂ 8 %), litten an schweren Gleichgewichtsproblemen, Schwindel oder Morbus Menière.
- Von den Patienten mit inneren Erkrankungen konnten insbesondere Personen mit Herzinsuffizienz, Herzklappenfehler und koronarer Herzkrankheit den Parcours nicht komplett durchlaufen (♀ 13 %, ♂ 16 %).
- Neurologie-Patienten (♀ 17 %, ♂ 5 %) mit einem kurzzeitig zurückliegenden Schlaganfall, Schädelhirntrauma, Epilepsie, Nervenschädigungen, Multiple Sklerose und Gehirntumor hatten ebenfalls deutliche Beeinträchtigungen beim Test.
- Orthopädische Patienten mit Bandscheibenvorfällen, Osteoporose und größeren rheumatischen sowie arthrotischen Beschwerden (♀ 19 %, ♂ 3 %) ließen ebenfalls mindestens ein Item aus.

- Bei den weiblichen gesunden Personen (5 %) waren die Gründe für die Nichtabsolvierung nicht ersichtlich.

Beim klassischen TUG benötigten diese Probanden mit im Mittel 7,6 s etwas mehr Zeit als der aktuell berechnete Mittelwert (6,5 s; s. Tab. 21, S. 81), lagen damit aber immer noch deutlich unterhalb des von Podsiadlo und Richardson (1991) bestimmten Trennpunktes (10 s).

Die Wahrscheinlichkeit alle Items zu lösen, nimmt bei dieser Untersuchung mit zunehmendem Alter um insgesamt 12 Prozent ab (Tab. 25).

*Tabelle 25: Lösungshäufigkeiten der Items bei den jeweiligen Altersquartilen (%; n = 764)*

<b>Alter (Jahre)</b>	<b>Alle Items erfüllt (%)</b>
≤ 49	92
50 - 57	90
58 - 63	90
> 63	80

Alle Teilnehmer konnten die Items „Aufstehen (Bank), Gehen“ und „Hindernis überwinden“ erfolgreich absolvieren. Obwohl nicht jeder Proband die restlichen Items löste, ergaben sich insgesamt niedrige Schwierigkeiten (Tab. 26, Spalten 3 - 5).

*Tabelle 26: Lösungshäufigkeiten der Items, differenziert nach Personengruppen, Alter und Geschlecht (%; n = 764)*

Personengruppe (n)	Aufstehen (Bank), Gehen	Hindernis überwinden	Kasten umstoßen	Hinsetzen/ Aufstehen (Boden)	Kasten tragen, Hinsetzen (Bank)
<b>Neurologie</b> (169)	100	100	95	98	94
<b>HNO</b> (224)	100	100	94	91	90
<b>Innere</b> (128)	100	100	92	98	89
<b>Orthopädie</b> (79)	100	100	86	96	93
<b>„Gesunde“</b> (153)	100	100	99	100	97
<b>Sportstudierende</b> (11)	100	100	100	100	100
<b>Altersquartile</b> (Jahre; n)					
≤ 49 (190)	100	100	97	97	94
50 - 57 (190)	100	100	96	97	94
58 - 63 (183)	100	100	97	96	95
> 63 (201)	100	100	89	94	90
<b>Geschlecht</b> (n)					
♀ (361)	100	100	98	96	97
♂ (403)	100	100	91	95	89

Neben der Lösungshäufigkeit wurden auch der erreichte Punktwert der Ausführungsbeurteilung (Bewertungsbereich 0 - 11; vgl. Tab. 14, S. 66) berechnet und Minimum-/Maximumwerte in Abhängigkeit der Personengruppe, Altersquartile und des Geschlechts ermittelt (Tab. 27).

*Tabelle 27: Testindex und Minimum-, Maximumwerte (Bewertungsbereich 0 - 11), in Abhängigkeit der Personengruppe, Altersquartile und des Geschlechts (n = 764)*

	Testindex (Punkte; M ± SD)	Minimum	Maximum
<b>Personengruppen</b> (n)			
Neurologie (169)	10 ± 1	7	11
HNO (224)	10 ± 1	5	11
Innere (128)	10 ± 1	6	11
Orthopädie (79)	9 ± 1	4	11
„Gesunde“ (153)	10 ± 1	7	11
Sportstudierende (11)	11 ± 0	11	11
<b>Altersquartile</b> (Jahre; n)			
≤ 49 (190)	10 ± 1	4	11
50 - 57 (190)	10 ± 1	6	11
58 - 63 (183)	10 ± 1	5	11
> 63 (201)	9 ± 1	6	11
<b>Geschlecht</b> (n)			
♀ (361)	9 ± 1	4	11
♂ (403)	10 ± 1	6	11

### 5.3.5.2 Testzeiten

#### *Personengruppe und Geschlecht*

Von 670 Personen, die alle Items lösten, konnten die Zeiten analysiert werden. Sie benötigten im Mittel  $22,1 \pm 9,0$  s für den Parcours.

Werden aus dem Gesamtzeitbereich (Minimum 7 bis Maximum 70 s) Quartile gebildet ( $\leq 16$  s, 17 - 20 s, 21 - 27 s,  $> 27$  s), zeigte sich, dass schnelle Testzeiten eher von den Gruppen der Gesunden und HNO-Patienten erzielt wurden. Die Patientengruppen Neurologie, Innere und Orthopädie benötigten meist höhere Testzeiten (Abb. 27). Für weitere Berechnungen werden dementsprechend die Gesunden und HNO-Patienten gemeinsam als „Körperlich Gesunde“ und die Neurologie-, Innere- und Orthopädiepatienten als „Körperlich Kranke“ betrachtet.

Lediglich bei der Bestimmung eines Grenzwertes beim TUG<sub>plus</sub> (Kap. 5.7.2) werden die HNO-Patienten, die zusätzlich die Diagnose Schwindel aufweisen, zu der Gruppe der „Körperlich Kranken“ hinzugefügt.

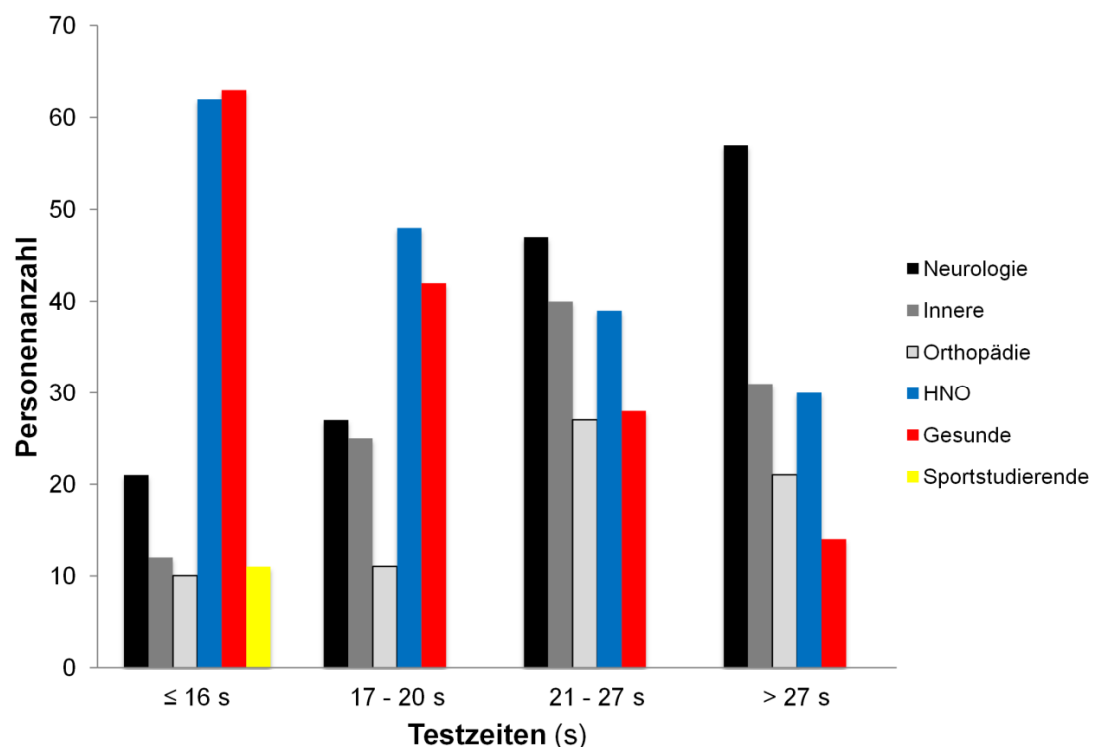


Abbildung 27: Verteilung der Personen- und Patientengruppen (Absolutwerte) in Abhängigkeit der Testzeiten (s) beim TUG<sub>plus</sub> ( $n = 670$ )

Bei „körperlich Kranken“ nimmt der Anteil an Personen mit zunehmender Testzeit zu. Bei „körperlich Gesunden“ nimmt hingegen der Anteil mit zunehmender Testzeit ab.

Tabelle 28 zeigt eine Differenzierung nach Personengruppe und geschlechtsabhängige Unterschiede.

*Tabelle 28: Testzeiten (s) der Teilnehmer, die alle Items absolvieren konnten und geschlechtsabhängige Unterschiede ( $M \pm SD$ ; MD = Median bei Nicht-Normalverteilung;  $n = 670$ )*

Personengruppe (n)	Geschlecht (n)	Zeit (s)	p-Wert
Neurologie (152)	♀ (59)	28,6 ± 8,6	0,000
	♂ (93)	23,0 ± 7,1	
HNO (181)	♀ (87)	23,3 ± 9,7	0,000
	♂ (94)	18,5 ± 8,9	
Innere (109)	♀ (58)	25,4 ± 7,7	0,025
	♂ (51)	22,3 ± 6,8	
Orthopädie (69)	♀ (39)	28,6 ± 9,8	0,020
	♂ (30)	22,8 ± 10,4	
„Gesunde“ (148)	♀ (92)	20,1 ± 7,1	0,000
	♂ (56)	15,7 ± 5,2	
Sportstudierende (11)	♀ (4)	10,2 ± 2,1; MD = 9,7	0,073
	♂ (7)	8,2 ± 0,7; MD = 8,4	

Weibliche Teilnehmer benötigten durchschnittlich etwa 20 % mehr Zeit als Männer (♀: 24,2 ± 9,2 s,  $n = 339$ ; ♂: 20,0 ± 8,3 s;  $p \leq 0,001$ ;  $d = 0,5$ ;  $n = 331$ ). In Abhängigkeit der Personengruppen waren Frauen mit neurologischen und/oder orthopädischen Erkrankungen durchschnittlich am langsamsten (28,6 s). Männliche Sportstudenten und „Gesunde“ waren am schnellsten.

Bei den Personen- bzw. Patientengruppen ergaben sich ebenfalls unterschiedliche Testzeiten ( $\eta^2 = 0,13$ ;  $F = 20,4$ ). Sportstudierende durchliefen den Parcours mit deutlichem Abstand am schnellsten ( $p \leq 0,001$ ). „Gesunde“ waren etwa um 10 - 40 % schneller ( $p \leq 0,001$ ) als Patienten mit neurologischen ( $d = 0,90$ ), inneren ( $d = 0,79$ ) und orthopädischen Erkrankungen ( $d = 0,87$ ). HNO-Patienten benötigten weniger Zeit als Patienten mit inneren ( $p \leq 0,001$ ;  $d = 0,37$ ) und orthopädischen Erkrankungen ( $p \leq 0,01$ ;  $d = 0,52$ ). Diese beiden Patientengruppen waren am langsamsten, unterschieden sich allerdings untereinander nicht (Abb. 28).

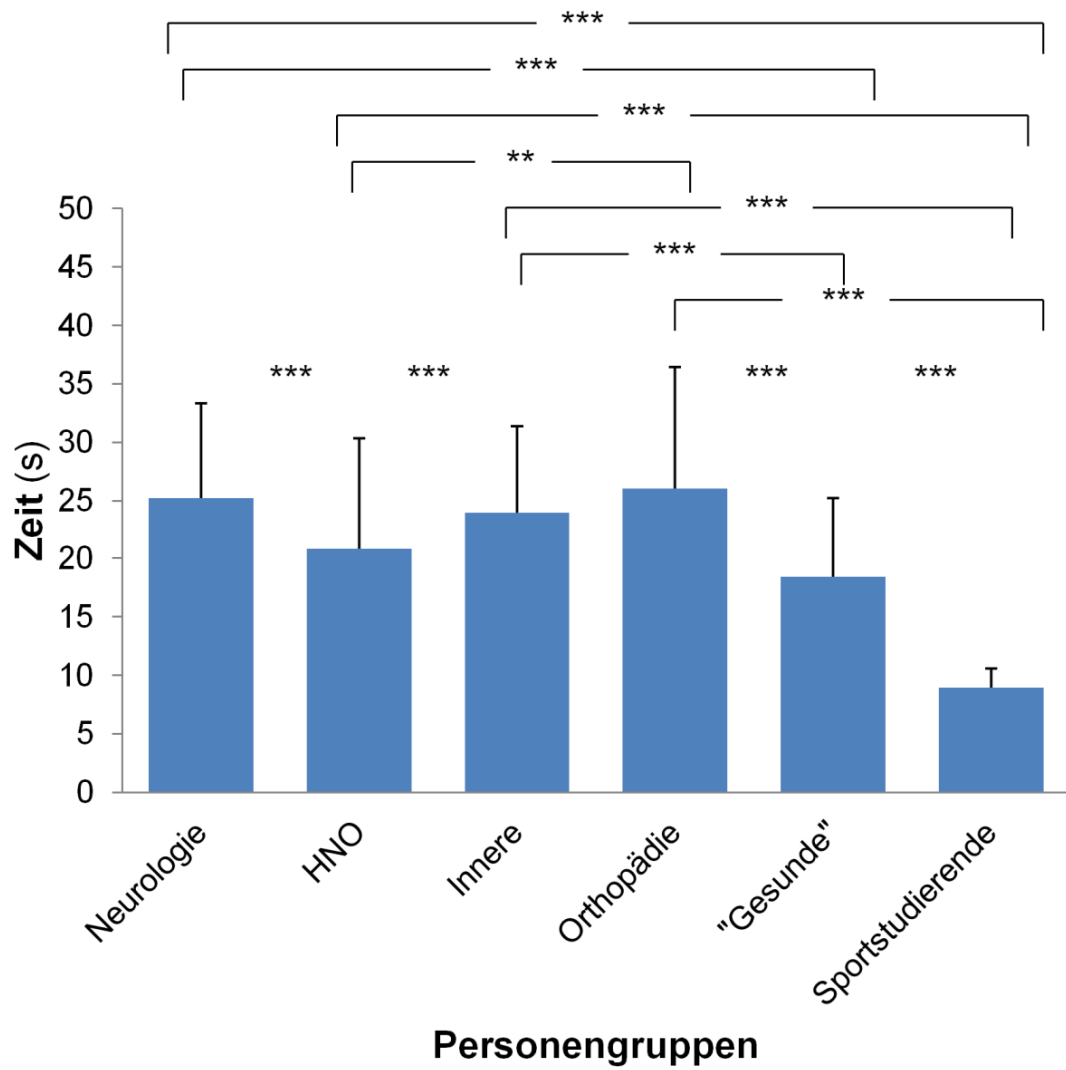


Abbildung 28: Testzeiten (s) des TUG<sub>plus</sub> bei unterschiedlichen Personengruppen (\*\*\*) =  $p \leq 0,001$ ; \*\* =  $p \leq 0,01$ ;  $n = 670$ )



### Alter

Aufgrund der geringen Stichprobengröße bei den Sportstudierenden erfolgt in den weiteren Darstellungen keine Ergebnisdifferenzierung. Die Gesamtstichprobe verringert sich somit auf  $n = 659$  (Tab. 29).

Die jüngste Probandin war 20 Jahre alt, litt an einer neurologischen Erkrankung und benötigte 19,1 s beim TUG<sub>plus</sub>. Der älteste Proband war 85 Jahre, Herz-Kreislaufpatient und benötigte 38 s für den Parcours. Insgesamt zeigte sich eine mit zunehmendem Alter ansteigende Testzeit (Tab. 29).

*Tabelle 29: Altersspezifische Testzeiten (s) von Teilnehmern, die alle Items absolvieren konnten, differenziert nach der Stichprobe ( $n = 659$ )*

Stichprobe (n)	Altersquartile (Jahre)	Zeit (s)
Gesamtstichprobe (659)	$\leq 49$	$18,4 \pm 7,9$
	50 - 57	$21,3 \pm 8,3$
	58 - 63	$23,0 \pm 8,3$
	$> 63$	$26,2 \pm 9,0$
	gesamt	$23,0 \pm 8,8$
Neurologie (152)	$\leq 48$	$20,2 \pm 6,0$
	49 - 53	$24,4 \pm 10,3$
	54 - 60	$24,4 \pm 8,6$
	$> 60$	$31,1 \pm 10,2$
	gesamt	$25,2 \pm 8,1$
HNO (181)	$\leq 49$	$19,4 \pm 8,3$
	50 - 57	$20,9 \pm 7,3$
	58 - 64	$22,2 \pm 10,9$
	$> 64$	$30,7 \pm 10,6$
	gesamt	$20,7 \pm 9,6$
Innere (109)	$\leq 51$	$21,5 \pm 7,3$
	52 - 58	$22,6 \pm 5,9$
	59 - 63	$24,1 \pm 6,5$
	$> 63$	$26,3 \pm 7,2$
	gesamt	$25,6 \pm 7,4$
Orthopädie (69)	$\leq 56$	$23,8 \pm 7,6$
	57 - 61	$22,4 \pm 7,2$
	62 - 66	$24,9 \pm 6,8$
	$> 66$	$31,4 \pm 10,7$
	gesamt	$26,8 \pm 10,3$
Gesunde (148)	$\leq 51$	$16,8 \pm 6,2$
	52 - 59	$19,0 \pm 6,0$
	60 - 65	$19,2 \pm 6,3$
	$> 65$	$25,6 \pm 9,7$
	gesamt	$18,9 \pm 5,9$

Bei der Gesamtstichprobe unterschieden sich die Zeiten in Abhängigkeit der Altersquartile, außer zwischen den Altersspannen 49 bis 57 und 57 bis 63 Jahren ( $p \leq 0,01$ ; Abb. 29).

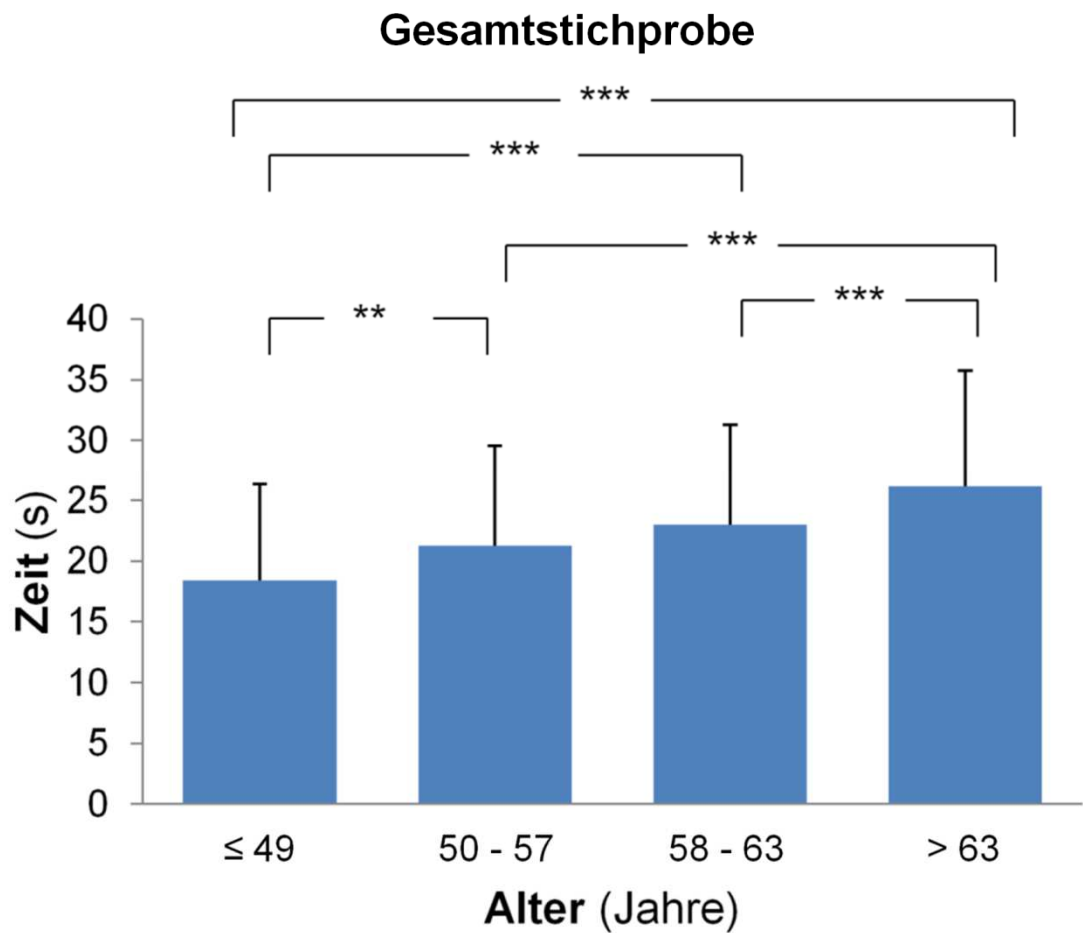


Abbildung 29: Altersspezifische Unterschiede der Testzeiten (s) beim  $TUG_{plus}$  der Gesamtstichprobe (\*\*\*) =  $p \leq 0,001$ ; \*\* =  $p \leq 0,01$ ;  $n = 659$ )

Bei neurologischen Patienten ergaben sich Unterschiede zwischen den Gruppen ( $\eta^2 = 0,14$ ;  $F = 6,8$ ; Abb. 30), insbesondere zwischen den Altersspannen  $\leq 48$  Jahren (20 s) und  $> 60$  Jahren (31 s;  $p \leq 0,01$ ;  $d = 1,1$ ).

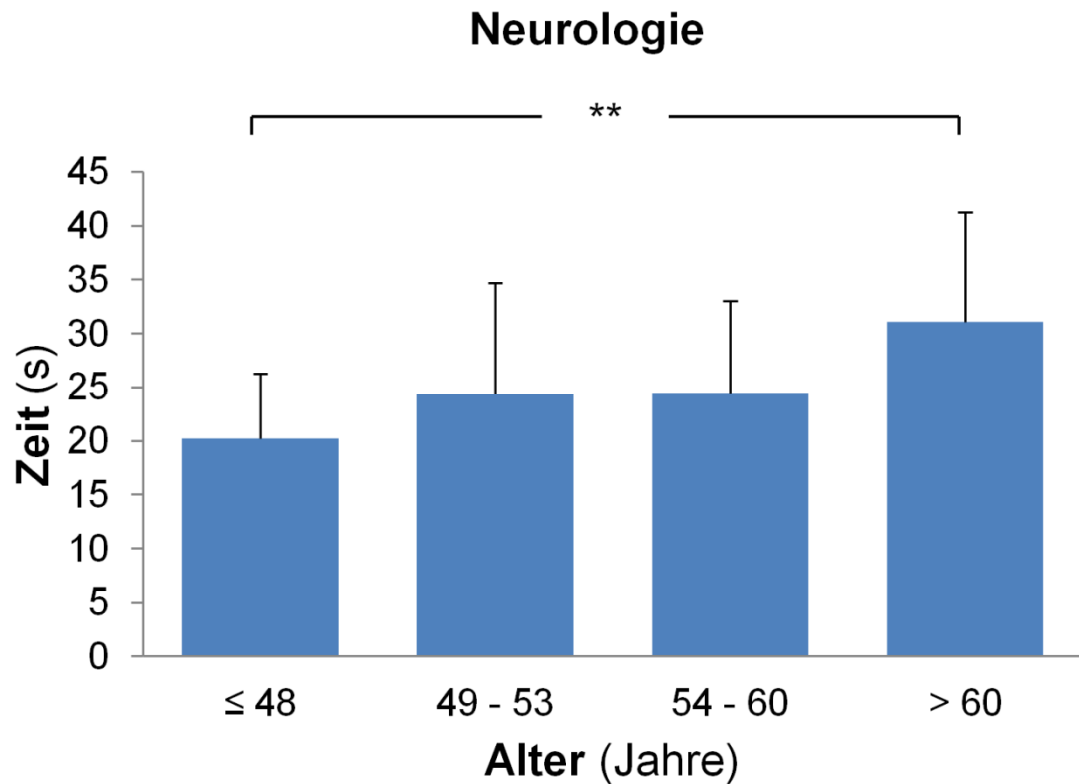


Abbildung 30: Neurologie – Altersabhängige Testzeiten (s) beim  $TUG_{plus}$  (\*\* =  $p \leq 0,01$ ;  $n = 152$ )

Bei HNO-Patienten konnten ebenfalls altersabhängige Testzeiten gemessen werden ( $\eta^2 = 0,11$ ;  $F = 7,7$ ; Abb. 31). Die Testzeiten von Probanden jünger als 64 Jahre unterschieden sich von denen bei Personen älter als 64 Jahre ( $\leq 49$  und  $> 64$  Jahre:  $p \leq 0,001$ ;  $d = 1,2$ ; 50 - 57 und  $> 64$  Jahre:  $p \leq 0,01$ ;  $d = 1,1$ ; 58 - 64 und  $> 64$  Jahre:  $p \leq 0,05$ ;  $d = 0,8$ ).

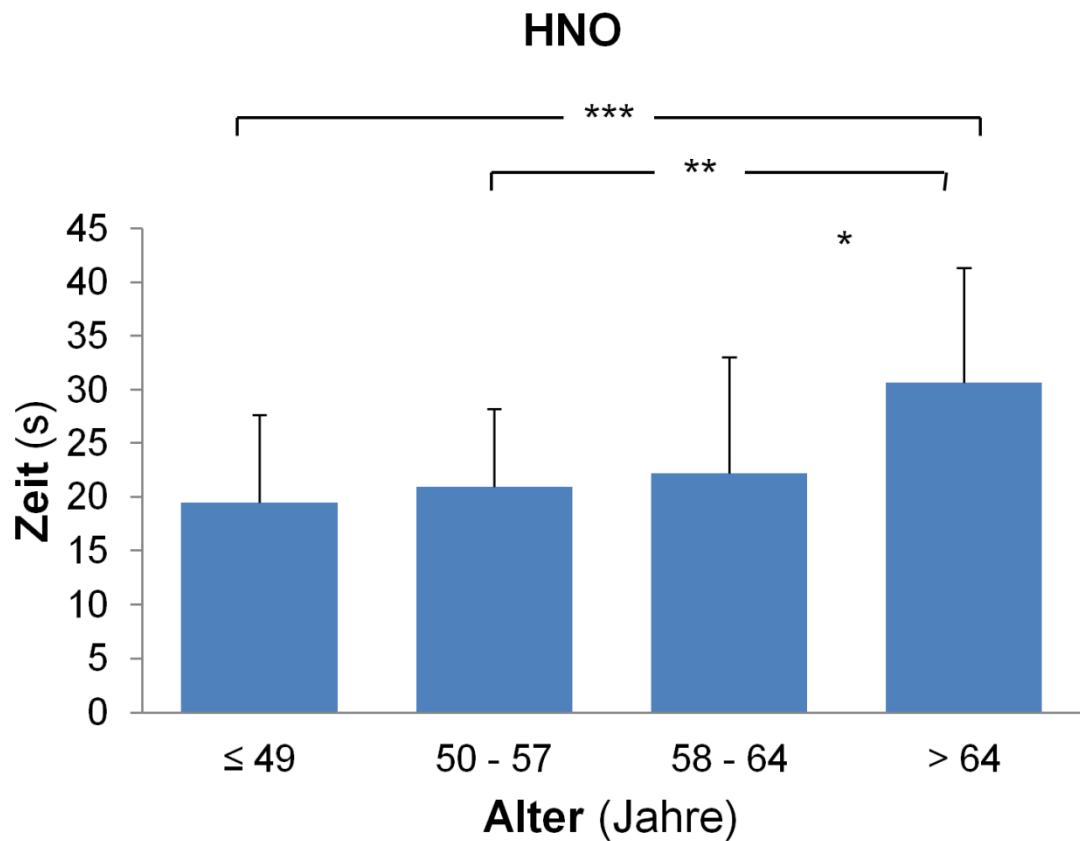


Abbildung 31: HNO – Altersabhängige Testzeiten (s) beim  $TUG_{plus}$  (\*\* =  $p \leq 0,01$ ;  $n = 181$ )

Die Testzeiten von Patienten mit inneren Erkrankungen stieg im Mittel zwar an, jedoch ergaben sich keine relevanten Unterschiede ( $F = 2,1$ ; Abb. 32).

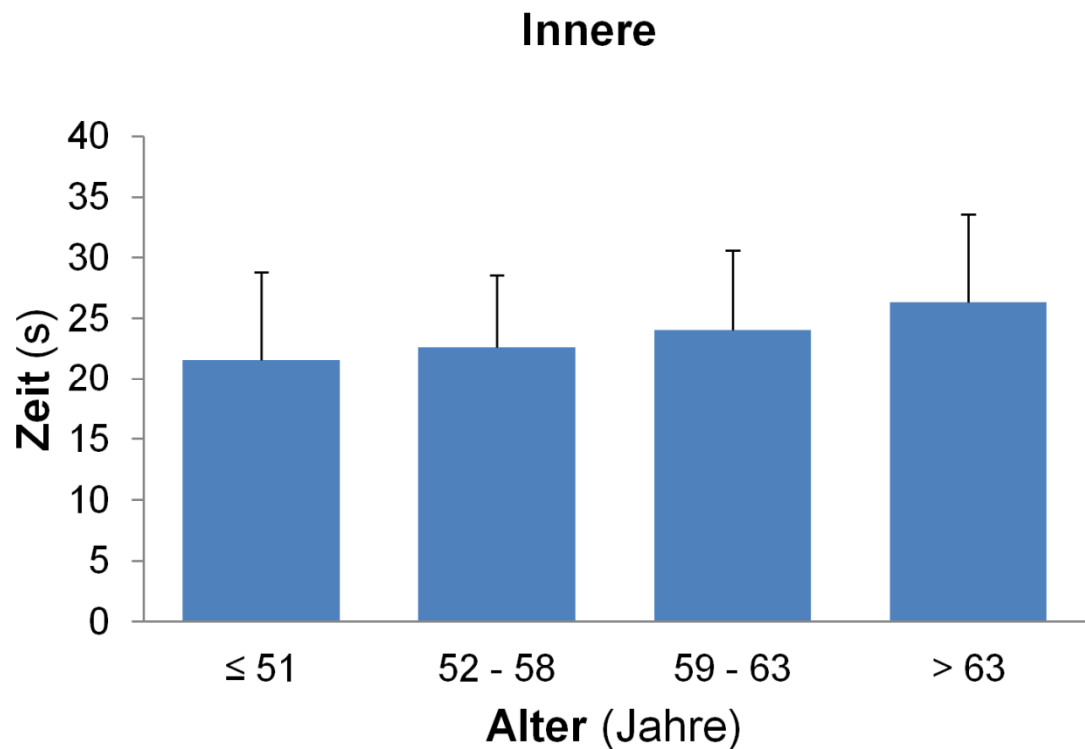


Abbildung 32: Innere – Altersabhängige Testzeiten (s) beim  $TUG_{plus}$  ( $n = 109$ )

Bei Orthopädie-Patienten ergaben sich altersabhängige Unterschiede hinsichtlich der Testzeiten ( $\eta^2 = 0,16$ ;  $F = 3,9$ ; Abb. 33). Personen im Alter zwischen 57 und 61 Jahren (22 s) waren schneller als Probanden über 66 Jahren (31 s;  $p \leq 0,05$ ;  $d = 1,0$ ).

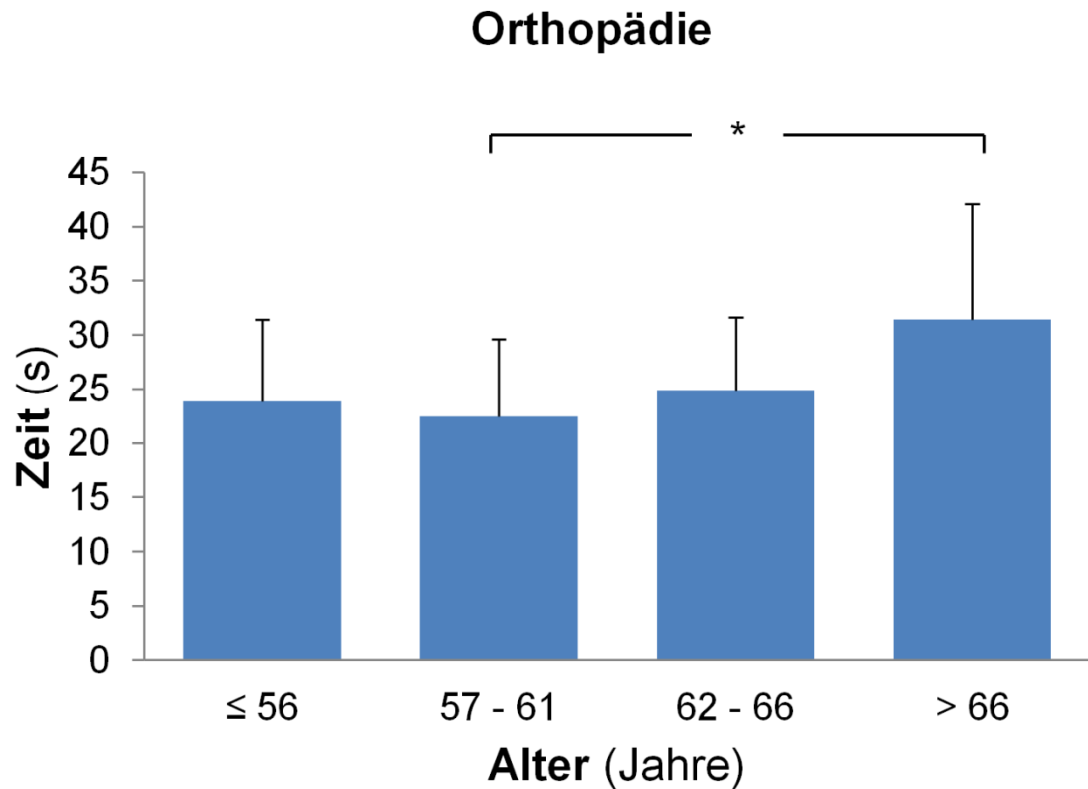


Abbildung 33: Orthopädie – Altersabhängige Testzeiten (s) beim  $TUG_{plus}$  (\* =  $p \leq 0,05$ ;  $n = 69$ )

Altersabhängige Unterschiede hinsichtlich der Testzeiten zeigten sich auch bei Gesunden ( $\eta^2 = 0,17$ ;  $F = 7,4$ ; Abb. 34). Die Geschwindigkeiten wurden jedoch erst ab einem Alter von 65 Jahren niedriger ( $\leq 51$  Jahre und  $> 65$  Jahre:  $p \leq 0,01$ ;  $d = 1,1$ ; 52 - 59 Jahre und  $> 65$  Jahre:  $p \leq 0,01$ ;  $d = 0,8$ ; 60 - 65 Jahre und  $> 65$  Jahre:  $p \leq 0,01$ ;  $d = 0,8$ ).

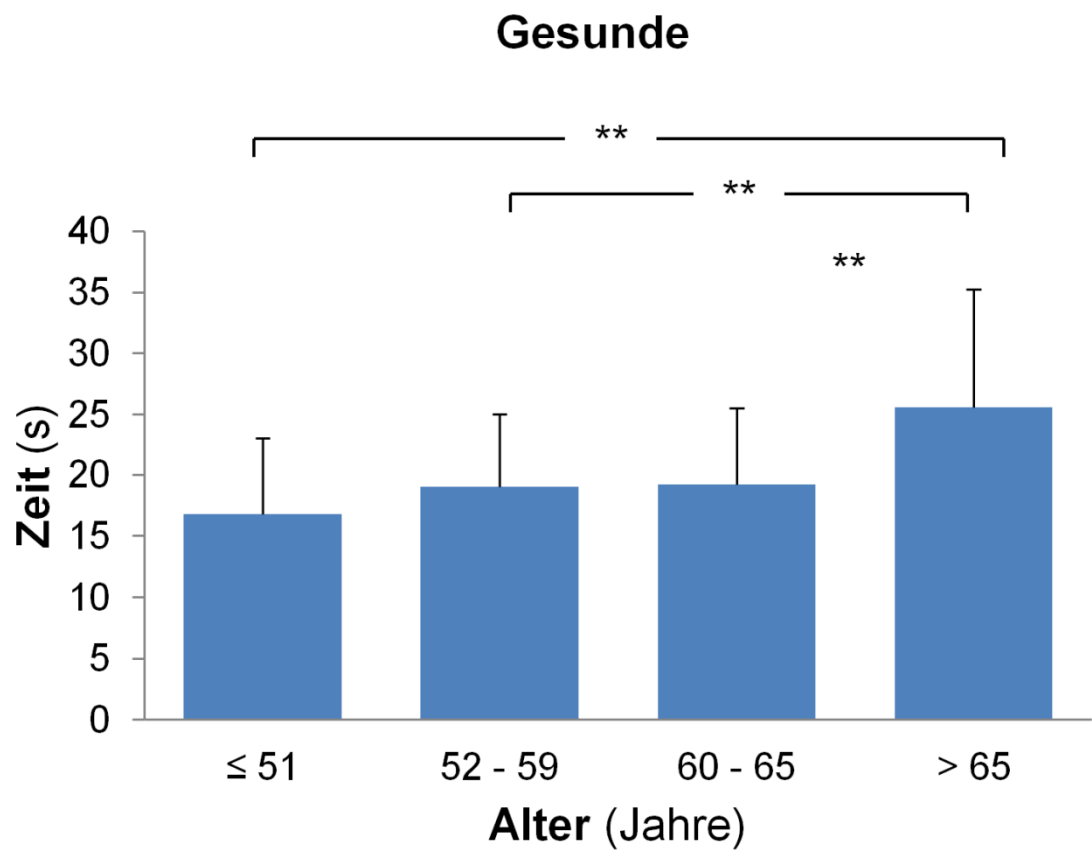


Abbildung 34: „Gesunde“ – Altersabhängige Testzeiten (s) beim  $TUG_{plus}$  (\*\* =  $p \leq 0,01$ ;  $n = 148$ )

### 5.3.5.3 Testzeiten, Ausführungsökonomie, Extremwert- bzw. Einzelfallbetrachtungen

#### *Testzeiten und Ausführungsökonomie*

Hinsichtlich der Bewegungsausführung erreichten die Probanden einen durchschnittlichen Testindex von  $9,7 \pm 1,1$  Punkte. Es zeigte sich eine mittlere Korrelation zwischen der Bewegungsausführung bzw. Ausführungsökonomie und der Testzeit ( $r = -0,63$ ;  $p \leq 0,001$ ; Abb. 35).

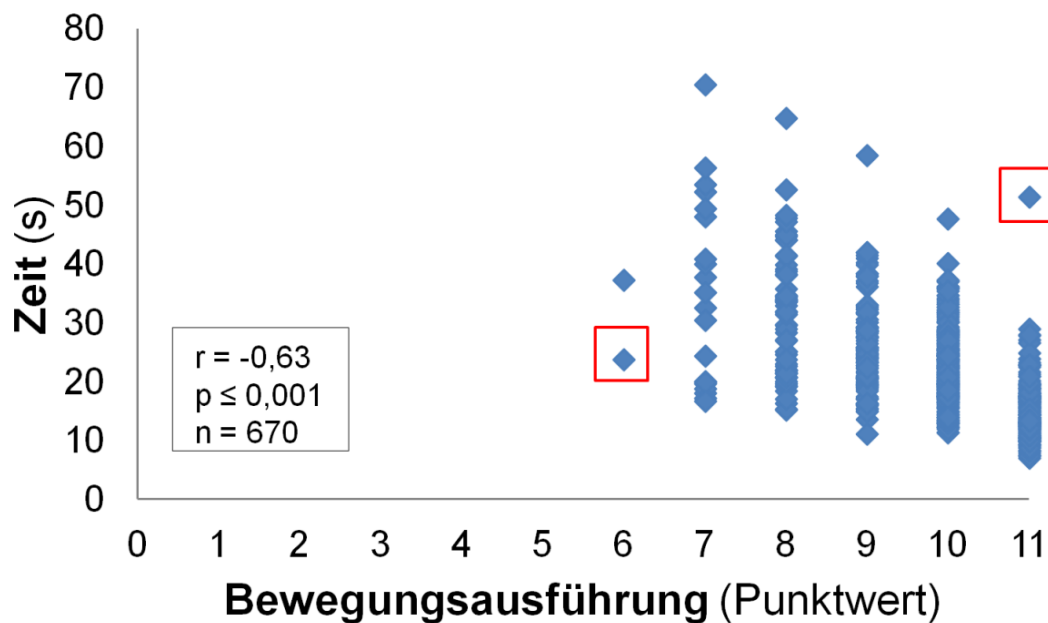


Abbildung 35: Zusammenhang ( $r$ ) zwischen Ausführungsökonomie (Punktwert) und Testzeit (s) beim TUG<sub>plus</sub> ( $n = 670$ ; Einzelwert =  )

Auffällige Ergebnisse sind bei zwei Personen zu erkennen (Einzeltwerte, s. Abb. 35). Eine Probandin (Alter: 75 Jahre, Brustkrebspatientin, Körpergröße: 163 cm, Körpergewicht: 55 kg) erzielte trotz maximal ökonomischer Bewegungsausführung (Punktwert 11) eine recht hohe Testzeit (55,5 s). Beim klassischen TUG benötigte sie allerdings auch nur 6,6 s. Eine zweite Probandin (Alter: 68 Jahre, Körpergröße: 166 cm, Körpergewicht: 60 kg, Diabetikerin) erzielte trotz mäßiger Bewegungsausführung (Punktwert 6) eine etwas oberhalb dem Mittel von rund 22 s liegende Testzeit (23,8 s). Sie konnte das Hindernis nur mit Mühe überwinden und benötigte zwei Versuche zum Umkippen des Kastens. Bei der Aufgabe den Kasten zu tragen, konnte sie lediglich 6 kg bewältigen, sodass sie die Strecke dreimal absolvieren musste. Beim klassischen TUG benötigte sie lediglich 5,9 s.



### Extremwerte

Um zu überprüfen wie die höchsten Testzeiten zustande kommen, wurden die zehn Probanden, die mehr als doppelt so lange als der Durchschnitt (s. Kap. 5.3.5.2, S. 96) zur Absolvierung des Parcours benötigten ( $> 48$  s), ausgewählt. Unter diesen Personen waren Neurologie-, HNO- und Innere-Patienten vertreten (Tab. 30).

*Tabelle 30: Extremwerte – Patienten mit den langsamsten Testzeiten (s) beim TUG<sub>plus</sub> ( $\geq 48$  s): anthropometrische Daten, Diagnosen, Testzeiten (s) beim klassischen TUG und TUG<sub>plus</sub> inkl. Ausführungsqualität (Punktwert, max. 11 Punkte)*

Gruppe (n)	Geschlecht	Alter (Jahre)	BMI (kg/m <sup>2</sup> )	Diagnose	TUG (s)	TUG <sub>plus</sub> (s; Punkte)
Neurologie (2)	A: ♀	52	26	Polyneuropathie	8,8	52,3; 7
	B: ♂	52	24	Spinalkanalstenose	10,1	52,7; 8
HNO (4)	A: ♀	57	22	Tinnitus, Schwindel	8,4	48,4; 8
	B: ♀	57	33	Schwindel, Schwerhörigkeit	7,8	53,6; 7
	C: ♂	66	33	Schwindel	10,0	56,4; 7
	D: ♀	63	28	Schwindel	8,5	70,6; 7
Innere (4)	A: ♀	75	21	Zustand nach Herzinfarkt	6,6	51,5; 10
	B: ♀	76	22	Hypertonie, Zustand nach Stent-Implantat	6,1	49,5; 7
	C: ♂	82	27	Zustand nach Herzinfarkt	10,8	58,5; 9
	D: ♀	83	24	Diabetes Typ 2, Hypertonie	8,4	64,9; 8

Bei HNO-Patienten sind die Einzelbefunde konträr zu den Ergebnissen der Gesamtstichprobe – sie waren beim TUG<sub>plus</sub> die schnellste Patientengruppe (s. Abb. 28, S. 98).

Eine HNO-Patientin (D) benötigte mit 70,6 s insgesamt mit Abstand die meiste Zeit und lag mit 8,5 s beim TUG auch oberhalb des aktualisierten (6,5 s), aber immer noch unterhalb des ursprünglichen Grenzwertes (10 s).

Bei der weiteren Evaluierung wurde aufgrund des recht hohen Zusammenhangs zwischen Testzeit und Bewegungsausführung die Beurteilung der Bewegungsausführung vernachlässigt, um die Ökonomie des TUG<sub>plus</sub> zu erhöhen.

## 5.4 Reliabilität

### 5.4.1 Interne Konsistenz

#### 5.4.1.1 Fragestellungen

Die interne Konsistenz eines Tests gibt Aufschluss darüber, in welchem Ausmaß die einzelnen Items in Beziehung stehen (Lienert, & Raatz, 1998, S. 201). Ein häufig verwendetes Maß ist hierbei Cronbachs Alpha ( $\alpha$ ). Die Höhe dieses Koeffizienten ist von den Korrelationen der Items abhängig.

Im Folgenden sollte die Frage beantwortet werden, inwiefern die Items des Screenings in Beziehung zueinander stehen und wie gut die Trennschärfe zu bewerten ist. Dies sollte geschlechtsunabhängig und -abhängig überprüft werden.

#### 5.4.1.2 Methodik

##### 5.4.1.2.1 Personenstichprobe

An der Untersuchung nahmen 166 Personen mit unterschiedlichen Gesundheitszuständen (Neurologie, HNO, Innere Medizin, Orthopädie, „Körperlich Gesunde“) teil (Tab. 31).

*Tabelle 31: Interne Konsistenz: Anthropometrische Daten der Personen- bzw. Patientengruppen ( $M \pm SD$ ;  $n = 166$ )*

Personengruppe (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Neurologie (45)	♀ (19)	51 $\pm$ 6	167 $\pm$ 7	75 $\pm$ 17
	♂ (26)	54 $\pm$ 8	177 $\pm$ 7	88 $\pm$ 16
HNO (59)	♀ (32)	50 $\pm$ 9	167 $\pm$ 7	69 $\pm$ 11
	♂ (27)	53 $\pm$ 7	180 $\pm$ 7	91 $\pm$ 10
Innere Medizin (25)	♀ (11)	60 $\pm$ 6	164 $\pm$ 4	67 $\pm$ 8
	♂ (14)	67 $\pm$ 8	173 $\pm$ 5	81 $\pm$ 10
Orthopädie (26)	♀ (16)	51 $\pm$ 16	165 $\pm$ 5	71 $\pm$ 12
	♂ (10)	51 $\pm$ 15	178 $\pm$ 7	87 $\pm$ 20
„Gesunde“ (11)	♀ (8)	61 $\pm$ 7	164 $\pm$ 7	63 $\pm$ 10
	♂ (3)	55 $\pm$ 5	185 $\pm$ 4	83 $\pm$ 7

#### 5.4.1.2.2 Variablenstichprobe

Der TUG<sub>plus</sub> wurde, wie in Kapitel 4.2.2 (S. 60) beschrieben, von allen Probanden durchgeführt. Als Parameter wurden die Zwischenzeiten der jeweiligen Items gemessen.

#### 5.4.1.2.3 Ablauf der Untersuchung

Die Probanden durchliefen den Parcours nach standardisierter Instruktion. Die Durchgänge wurden mittels Videokamera aufgezeichnet. Später konnten so die benötigten Zeiten (s) für jedes Item über die Software Dartfish gemessen werden (Abb. 36).

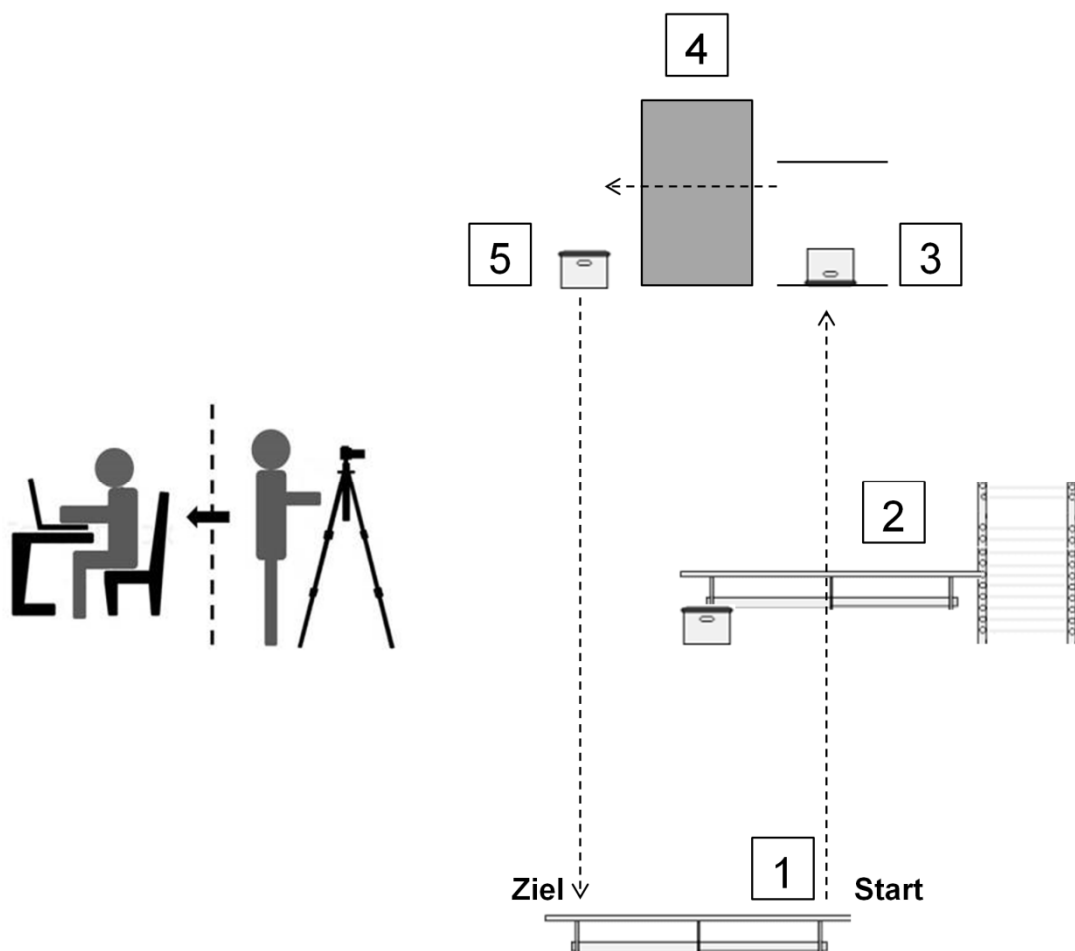


Abbildung 36: Untersuchungsdurchführung mit Itemmessung und -auswertung

Für die Zeitmessung der einzelnen Items war eine Festlegung exakter Start- und Endpunkte erforderlich (Ashford, & De Souza, 2000; Tab. 32).

*Tabelle 32: Start- und Endpunkte der Messung der Zwischenzeiten (s) bei den einzelnen Items*

Item	Start der Zeitmessung	Ende der Zeitmessung
Aufstehen (Bank), Gehen	Erster Startimpuls (Knie, Rücken, Fuß, ...)	Erste Berührung des Hindernisses
Hindernis überwinden	Berühren des Hindernisses	Erste Kastenberührung
Kasten umstoßen	Erste Kastenberührung	Kasten in Endposition
Hinsetzen/ Aufstehen (Boden)	Fuß berührt die Matte bzw. Fuß vor der Matte	Erste Kastenberührung
Kasten tragen, Hinsetzen (Bank)	Erste Kastenberührung	Erste Bankberührung

#### 5.4.1.3 Hypothesen

Hypothese 7: Die Items des TUG<sub>plus</sub> korrelieren untereinander (Messparameter Zeit), sodass eine akzeptable interne Konsistenz für Männer und Frauen, sowie geschlechtsunabhängig angenommen werden kann ( $r > 0,4$ ;  $\alpha \geq 0,7$ ).

Hypothese 8: Die Items zeigen bei Männer und Frauen sowie geschlechtsunabhängig Trennschärfekoeffizienten von mindestens  $r_{it} = 0,3$ , sodass jedes Item das Gesamtergebnis des TUG<sub>plus</sub> repräsentiert.

#### 5.4.1.4 Statistik

Die interne Konsistenz wurde mittels Inter-Itemkorrelationen bestimmt. Nach Logarithmisierung der Zeiten jedes Items, waren die Daten normalverteilt. Die Item-Statistik wurde genutzt, um die Trennschärfekoeffizienten ( $r_{it}$ ) zu beurteilen. Cronbachs alpha ( $\alpha$ ) diente als Maß der internen Konsistenz.

### 5.4.1.5 Ergebnisse

#### 5.4.1.5.1 Item Zeiten und Korrelationen

Tabelle 33 zeigt die mittleren Zeiten der einzelnen Items bei Männern, Frauen und geschlechtsunabhängig.

*Tabelle 33: Zwischenzeiten (s) der Items, geschlechtsabhängig und -unabhängig ( $M \pm SD$ ;  $n = 166$ )*

Item	Testzeit (s)		
	Männer ( $n = 80$ )	Frauen ( $n = 86$ )	Gesamt ( $n = 166$ )
Aufstehen (Bank), Gehen	$2,4 \pm 0,8$	$2,5 \pm 0,5$	$2,5 \pm 0,7$
Hindernis Überwinden	$3,3 \pm 1,4$	$3,8 \pm 1,2$	$3,6 \pm 1,3$
Kasten Umstoßen	$3,3 \pm 1,8$	$3,8 \pm 2,0$	$3,5 \pm 1,9$
Hinsetzen/ Aufstehen (Boden)	$6,9 \pm 2,5$	$7,0 \pm 2,3$	$7,0 \pm 2,4$
Kasten tragen, Hinsetzen (Bank)	$7,7 \pm 2,3$	$8,8 \pm 2,1$	$8,2 \pm 2,3$

Unabhängig vom Geschlecht zeigten alle Items mittlere Korrelationen ( $p \leq 0,01$ ) zueinander bei einem  $\alpha$  von 0,85 (Tab. 34).

*Tabelle 34: Geschlechtsunabhängige Itemkorrelationen ( $r$ ; \*\*\* =  $p \leq 0,001$ ;  $n = 166$ )*

	Aufstehen (Bank), Gehen	Hindernis überwinden	Kasten umstoßen	Hinsetzen/ Aufstehen (Boden)	Kasten tragen, Hinsetzen (Bank)
Aufstehen (Bank), Gehen	-	0,61***	0,55***	0,69***	0,65***
Hindernis überwinden	0,61***	-	0,56***	0,75***	0,68***
Kasten umstoßen	0,55***	0,56***	-	0,55***	0,59***
Hinsetzen/ Aufstehen (Boden)	0,69***	0,75***	0,55***	-	0,71***
Kasten tragen, Hinsetzen (Bank)	0,65***	0,68***	0,59***	0,71***	-

Bei Männern zeigten die Items mittlere bis hohe Korrelationen zueinander ( $p \leq 0,01$ ) bei einem  $\alpha$  von 0,89 (Tab. 35).

*Tabelle 35: Itemkorrelationen bei Männern ( $r$ ; \*\*\* =  $p \leq 0,001$ ;  $n = 80$ )*

	Aufstehen (Bank), Gehen	Hindernis überwinden	Kasten umstoßen	Hinsetzen/ Aufstehen (Boden)	Kasten tragen, Hinsetzen (Bank)
Aufstehen (Bank), Gehen	-	0,62***	0,67***	0,71***	0,80***
Hindernis überwinden	0,62***	-	0,65***	0,74***	0,75***
Kasten umstoßen	0,67***	0,65***	-	0,66***	0,74***
Hinsetzen/ Aufstehen (Boden)	0,71***	0,74***	0,66***	-	0,78***
Kasten tragen, Hinsetzen (Bank)	0,80***	0,75***	0,74***	0,78***	-

Bei Frauen zeigten die Items ebenfalls mittlere bis hohe Korrelationen zueinander ( $p \leq 0,01$ ) bei einem  $\alpha$  von 0,79 (Tab. 36).

*Tabelle 36: Itemkorrelationen bei Frauen ( $r$ ; \*\*\* =  $p \leq 0,001$ ;  $n = 86$ )*

	Aufstehen (Bank), Gehen	Hindernis überwinden	Kasten umstoßen	Hinsetzen/ Aufstehen (Boden)	Kasten tragen, Hinsetzen (Bank)
Aufstehen (Bank), Gehen	-	0,61***	0,40***	0,68***	0,44***
Hindernis überwinden	0,61***	-	0,44***	0,78***	0,54***
Kasten umstoßen	0,40***	0,44***	-	0,44***	0,40***
Hinsetzen/ Aufstehen (Boden)	0,68***	0,78***	0,44***	-	0,64***
Kasten tragen, Hinsetzen (Bank)	0,44***	0,54***	0,40***	0,64***	-

#### 5.4.1.5.2 Trennschärfe

Die Koeffizienten der Items lagen geschlechtsunabhängig zwischen  $r_{it} = 0,64$  (Kasten umstoßen) und  $0,79$  (Hinsetzen/Aufstehen – Boden). Durch Weglassen des Items „Aufstehen (Bank), Gehen“ ( $r_{it} = 0,74$ ) könnte die interne Konsistenz geringfügig erhöht werden (von  $\alpha = 0,85$  auf  $\alpha = 0,86$ ; Tab. 37).

*Tabelle 37: Geschlechtsunabhängige Statistik der jeweiligen Items (Messkriterium: Testzeit in s) in Bezug zur Testzusammensetzung (n = 166)*

Item (s)	Skalen- Mittelwert ohne jeweiliges Item	Skalen- varianz ohne jeweiliges Item	Trenn- schärfe- koeffizient	Cronbachs Alpha ohne jeweiliges Item
Aufstehen (Bank), Gehen	22,2	45,8	0,74	0,86
Hindernis überwinden	21,4	38,5	0,77	0,81
Kasten umstoßen	21,2	35,0	0,64	0,83
Hinsetzen/ Aufstehen (Boden)	17,7	27,3	0,79	0,80
Kasten tragen, Hinsetzen (Bank)	16,5	28,4	0,77	0,80

Bei Männern ergaben sich Werte von  $r_{it} = 0,75$  (Kasten umstoßen) bis  $0,87$  (Kasten tragen, Hinsetzen – Bank). Das Weglassen von Items führt hier zu keiner Verbesserung der internen Konsistenz (Tab. 38).

*Tabelle 38: Itemstatistik bei Männern (Messkriterium: Testzeit in s; n = 80)*

Item (s)	Skalen- mittelwert ohne jeweiliges Item	Skalen- varianz ohne jeweiliges Item	Trenn- schärfe- koeffizient	Cronbachs Alpha ohne jeweiliges Item
Aufstehen (Bank), Gehen	21,2	50,4	0,80	0,89
Hindernis überwinden	20,3	43,4	0,79	0,86
Kasten umstoßen	20,4	39,6	0,75	0,86
Hinsetzen/ Aufstehen (Boden)	16,8	30,4	0,81	0,86
Kasten tragen, Hinsetzen (Bank)	16,0	31,7	0,87	0,83

Bei Frauen zeigten sich geringere Item-Trennschärfen als bei Männern ( $r_{it} = 0,49$  Kasten umstoßen bis  $r_{it} = 0,77$  Hinsetzen/Aufstehen – Boden). Durch Weglassen des Items „Aufstehen (Bank), Gehen“ könnte die interne Konsistenz geringfügig erhöht werden (von  $\alpha = 0,79$  auf  $\alpha = 0,80$ ; Tab. 39).

*Tabelle 39: Itemstatistik bei Frauen (Messkriterium: Testzeit in s;  $n = 86$ )*

<b>Item (s)</b>	<b>Skalen- mittelwert ohne jeweiliges Item</b>	<b>Skalen- varianz ohne jeweiliges Item</b>	<b>Trenn- schärfe- koeffizient</b>	<b>Cronbachs Alpha ohne jeweiliges Item</b>
Aufstehen (Bank), Gehen	23,5	38,5	0,65	0,80
Hindernis überwinden	22,1	31,7	0,74	0,73
Kasten umstoßen	22,1	28,4	0,49	0,79
Hinsetzen/ Aufstehen (Boden)	18,9	21,8	0,77	0,68
Kasten tragen, Hinsetzen (Bank)	17,2	24,3	0,63	0,74



## 5.4.2 Paralleltest-Reliabilität

### 5.4.2.1 Fragestellungen

In dieser Untersuchung<sup>6</sup> sollten die Sequenzeffekte (Bortz, & Döring, 2005, S. 196) überprüft werden. Es galt zu untersuchen, ob die Reihenfolge der Items Einfluss auf das Ergebnis haben.

### 5.4.2.2 Methodik

#### 5.4.2.2.1 Personenstichprobe

An der Untersuchung nahmen 49 Probanden (14 ♀) teil, die eine stationäre Rehabilitation in der Bosenberg Klinik in St. Wendel durchliefen. Alle Patienten wurden in der Anreisewoche getestet (Tab. 40).

*Tabelle 40: Paralleltest-Reliabilität – Anthropometrische Daten der beiden Patientengruppen ( $M \pm SD$ ;  $n = 49$ )*

Gruppe	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Neurologie	♀ (9)	50 ± 5	167 ± 8	73 ± 17
	♂ (10)	52 ± 10	175 ± 5	87 ± 19
HNO	♀ (5)	52 ± 6	163 ± 11	63 ± 10
	♂ (5)	55 ± 5	181 ± 5	93 ± 12

#### 5.4.2.2.2 Variablenstichprobe

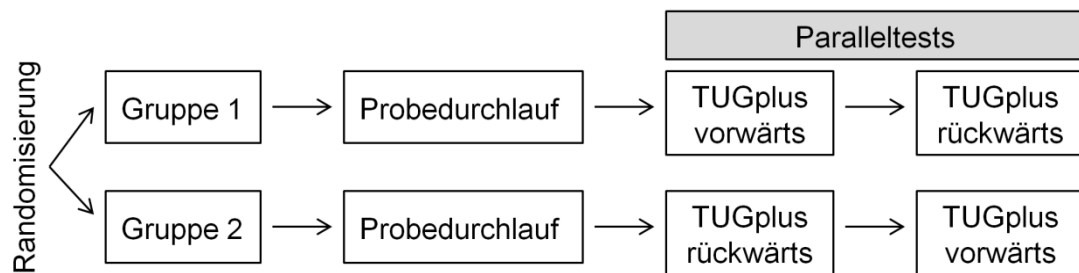
Beim TUG<sub>plus</sub> wurde die Geschwindigkeit (s) gemessen und der subjektive Anstrengungsgrad mittels Borg-Skala (6 bis 20) registriert (s. Tab. 15, S. 73).

#### 5.4.2.2.3 Ablauf der Untersuchung

Die Patienten wurden randomisiert in zwei Gruppen eingeteilt. Nach einem Probedurchlauf absolvierten beide Gruppen den Parcours in unterschiedlicher Itemreihenfolge jeweils zweimal hintereinander. Gruppe 1 führte den Test zunächst in vorgesehener Reihenfolge (vorwärts) und anschließend in umgekehrter Reihenfolge (rückwärts) durch. Gruppe 2 startete in umgekehrter Reihenfolge (Abb. 37 a, b).

<sup>6</sup> Die Daten wurden im Rahmen einer im Arbeitsbereich durchgeführten Masterarbeit in Zusammenarbeit mit Frau Thomas-Frank (2016) erhoben.

a)



b)

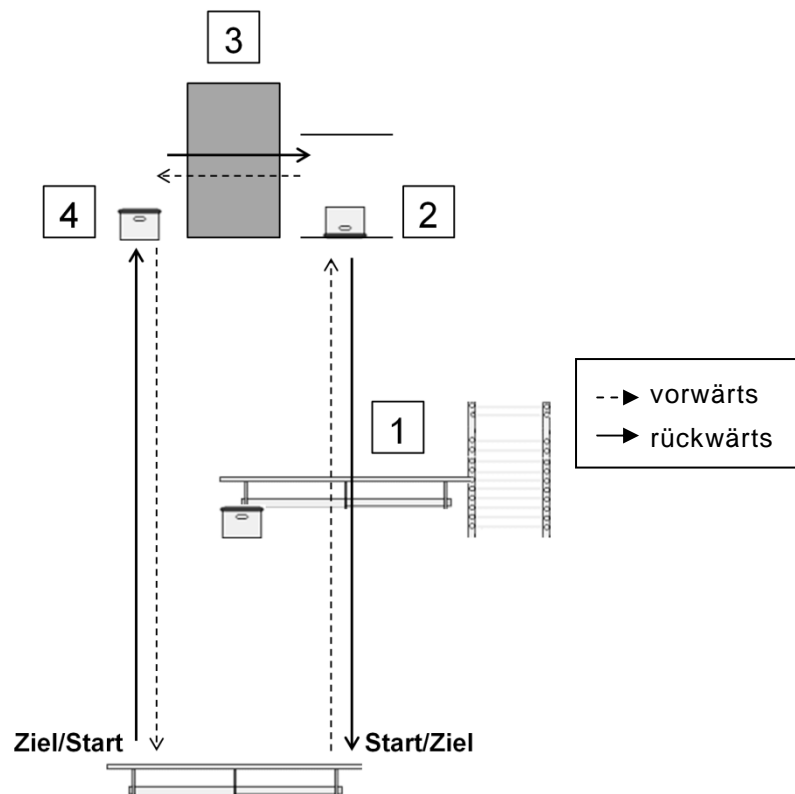


Abbildung 37: Paralleltest-Reliabilität – a) Untersuchungsdesign, b) Testablauf

#### 5.4.2.3 Hypothesen

- Hypothese 9: Die Testzeiten (s) der vorgesehenen und umgekehrten Itemreihenfolge korrelieren untereinander ( $r > 0,4$ ;  $p \leq 0,05$ ).
- Hypothese 10: Der subjektive Anstrengungsgrad (Borg-Skala 6 - 20) bei der vorgesehenen und umgekehrten Itemreihenfolge korreliert untereinander ( $r > 0,4$ ;  $p \leq 0,05$ ).

#### 5.4.2.4 Statistik

Die Testzeiten waren normal- und die Borg-Werte nicht normalverteilt. Zur Berechnung der Zusammenhänge der Testzeiten wurde die Produkt-Moment-Korrelation nach Pearson und bei den Borg-Werten der Rangkorrelationskoeffizient nach Spearman berechnet.

Zur Überprüfung des Einflusses der Kovariate Alter, BMI und Erkrankung auf die Testzeiten wurde eine multivariate Varianzanalyse durchgeführt.

#### 5.4.2.5 Ergebnisse

Tabelle 41 zeigt die durchschnittlich benötigten Testzeiten und Borg-Werte bei der ursprünglichen und umgekehrten Itemreihenfolge.

*Tabelle 41: Testzeiten (s; n = 29) und subjektiver Anstrengungsgrad (Borg-Wert; n = 24) in Abhängigkeit der Itemreihenfolge. Korrelationen der Variablen zwischen der klassischen und umgekehrten Reihenfolge ( $M \pm SD$ ;  $r$ ; \*\*\* =  $p \leq 0,001$ )*

Kriterium (n)	A: Ursprüngliche Reihenfolge	B: Umgekehrte Reihenfolge	Korrelation (r) A, B
Zeit (29)	24,2 $\pm$ 5,5	25,0 $\pm$ 5,3	0,87***
Borg-Wert (24)	10,3 $\pm$ 3,1	10,0 $\pm$ 3,2	0,86***

Es bestehen hohe Zusammenhänge zwischen den Testzeiten der ursprünglichen und der umgekehrten Parcours-Reihenfolge. Ebenfalls hohe Zusammenhänge können hinsichtlich des subjektiven Anstrengungsgrades beobachtet werden.

Die Testzeiten unterscheiden sich nicht in Abhängigkeit der Kovariate Alter, BMI und Erkrankungen ( $p > 0,05$ ). Somit sind die Testzeiten unabhängig davon, in welcher Reihenfolge der Parcours durchgeführt wird.

## 5.5 Responsivität

### 5.5.1 Fragestellungen

In dieser Untersuchung sollte die Frage beantwortet werden, ob anhand des TUG<sub>plus</sub> Änderungen auf motorischer Ebene nach einer zweiwöchigen, stationären Therapiephase aufgezeigt werden können<sup>7</sup>. Hierbei wird von Änderungssensitivität (Schuck, 2000) bzw. Responsivität (Guyatt, Townsend, Berman, & Keller, 1987) gesprochen. Diese wird bei Motoriktests häufig maßgeblich dadurch beeinflusst, ob der Test vorher geübt oder aus ökonomischen Gründen ungeübt durchgeführt wird. Daher galt es zunächst die Gewichtung von Testeffekten zu überprüfen.

#### 5.5.1.1 Methodik

##### 5.5.1.1.1 Personenstichprobe

Insgesamt nahmen 160 Probanden an den Testungen teil. Nach Ausschluss von Extremwerten und Dropouts gingen Daten von 126 Patienten (49 % Neurologie; 52 % HNO) in die Analyse ein (Tab. 42).

*Tabelle 42: Testeffekt und Responsivität – Anthropometrische Daten der getesteten Patientengruppen ( $M \pm SD$ ;  $n = 126$ )*

Überprüfung	Patientengruppe (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Testeffekt	HNO (36)	♀ (13)	45 ± 13	165 ± 5	72 ± 17
		♂ (23)	50 ± 12	180 ± 7	91 ± 13
	Neurologie (30)	♀ (9)	46 ± 15	168 ± 3	75 ± 15
		♂ (21)	57 ± 9	179 ± 6	86 ± 13
Responsivität	HNO (29)	♀ (9)	46 ± 13	165 ± 6	69 ± 15
		♂ (20)	51 ± 11	179 ± 6	92 ± 14
	Neurologie (31)	♀ (12)	50 ± 12	169 ± 4	71 ± 12
		♂ (19)	58 ± 7	177 ± 8	91 ± 15

##### 5.5.1.1.2 Variablenstichprobe

Alle Probanden absolvierten den TUG<sub>plus</sub> nach standardisierter Instruktion mindestens ein Mal (s. Kap. 4.2.2, S. 60).

<sup>7</sup> Die Daten wurden im Rahmen einer im Arbeitsbereich durchgeführten Examensarbeit (Backes, 2016) erhoben und die Ergebnisse im Abstract- und Tagungsband *Bewegung, Raum und Gesundheit* (Kaczmarek et al., 2016a; Kaczmarek et al., 2017b) und im Tagungsband *Prävention und Rehabilitation in Zeiten der Globalisierung* (Kaczmarek et al., 2017a) veröffentlicht.

### 5.5.1.1.3 Treatmentstichprobe

#### Zweiwöchige, stationäre Rehabilitation

Die Patienten durchliefen je nach Erkrankung unterschiedliche Therapien in der Bosenberg Klinik in St. Wendel (Saarland). Die Therapieformen variierten, auch innerhalb der Krankheitsbilder. Die folgende Auflistung beschränkt sich auf die physischen Therapieformen:

#### *Gesundheitssport*

In der Bosenberg Klinik durchläuft jeder Patient zu Beginn der Rehabilitation die Motorische Basisdiagnostik – ein Testverfahren zur Aufdeckung motorischer Defizite (Wydra, 1996). Für den Gesundheitssport geeignete Personen zeigen keine motorische Auffälligkeiten und können mindestens zwei Minuten am Stück laufen. Der Gesundheitssport zielt auf die Verbesserung der motorischen Leistungsfähigkeit ab und beinhaltet Ausdauer-, Kraft-, Koordinations- und Beweglichkeitstraining. Zudem besteht die Möglichkeit an einem Schwimmtraining teilzunehmen.

#### *Wirbelsäulentherapie*

Die Teilnahme setzt voraus, dass Personen in der Lage sind, sich selbstständig auf den Boden zu setzen und wieder aufzustehen. Inhalte sind Dehn- und Kräftigungsübungen als Bestandteil der Rückenschule.

#### *Wassergymnastik*

An dieser Trainingsform nehmen Patienten mit Übergewicht und orthopädischen Beschwerden teil. Das Training beinhaltet gymnastische Übungen im Wasser zur Verbesserung der Kraft, Beweglichkeit und Koordination.

#### *Herzsport*

Die Fähigkeit mindestens zwei Minuten am Stück gehen zu können, ist Teilnahmevoraussetzung. Das Ziel ist die Verbesserung konditioneller und koordinativer Fähigkeiten.

#### *Gehtraining*

Das Gehtraining ist ein zusätzliches Angebot zum Herzsport. Hierbei werden Gehstrecken von mindestens 1,5 km bewältigt.

### *Bewegungsbad*

Im Bewegungsbad werden leichte gymnastische Übungen zur muskulären Lockerung im Warmwasser (32°; 120 cm Wassertiefe) durchgeführt.

#### 5.5.1.1.4 Ablauf der Untersuchung

Die Teilnehmer wurden in zwei Gruppen eingeteilt. Zur Sicherstellung, dass sich die Grunderkrankungen gleichmäßig verteilen, erfolgte die Zuteilung blockrandomisiert (Abb. 39).

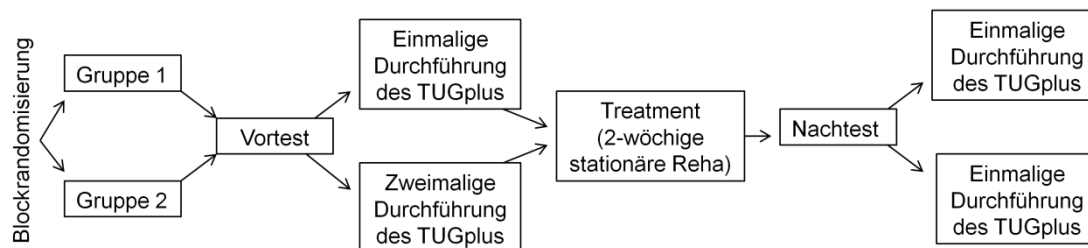


Abbildung 38: Untersuchungsdesign Testeffekt und Responsivität

### *Vortest*

Um Auswirkungen der Therapie bereits beim Vortest auszuschließen, wurde die Untersuchung am ersten Kliniktag nach Anreise der Patienten durchgeführt. Zur Überprüfung von Testeffekten absolvierten Gruppe 1 und Gruppe 2 einen Vortest ( $n = 66$ ). Gruppe 1 absolvierte den Parcours einmal, während Gruppe 2 diesen zweimal durchlief.

### *Treatment*

Alle Untersuchungsteilnehmer absolvierten ein zweiwöchiges, stationäres, sporttherapeutisches Rehabilitationsprogramm. Während diesem Zeitraum nahmen die Patienten an verschiedenen Trainingsprogrammen teil (s. Kap. 5.4.1.1.3, S. 119).

### *Nachtest*

Zur Beurteilung der Responsivität wurde ein Wiederholungstest durchgeführt ( $n = 60$ ). Hierzu wurden alle Patienten (Gruppe 1, 2) eingeschlossen, die am Vortest teilnahmen. Der Nachtest fand genau zwei Wochen nach dem Vortest statt. Um den Einfluss kognitiver Aspekte zu minimieren, wurde der TUG<sub>plus</sub> erneut verbal erklärt und demonstriert.

Alle Teilnehmer absolvierten den TUG<sub>plus</sub> einmal, unabhängig davon in welcher Gruppe sich die Patienten befanden.

#### 5.5.1.2 Hypothesen

Hypothese 11: Testeffekt: Die Testzeiten (s) von Patienten der Neurologie und HNO zeigen zwischen Vortest 1 und 2 eine mindestens mittlere Korrelation ( $r > 0,40$ ;  $p \leq 0,05$ ).

Hypothese 12: Responsivität: Die Patienten sind nach einem zweiwöchigen Therapieprogramm signifikant schneller beim TUG<sub>plus</sub> ( $p \leq 0,05$ ;  $d > 0,5$ ).

#### 5.5.1.3 Statistik

Da die Ergebnisse eine Normalverteilung aufwiesen, wurde zur Beurteilung von Testeffekten der Produkt-Moment-Korrelationskoeffizient nach Pearson verwendet. Nach zweiwöchiger Intervention erfolgte die Signifikanzprüfung aufgrund normalverteilter Daten mittels t-Test für abhängige Variablen. Die Effektstärke (Cohens d) wurde zur Untersuchung von Leistungsveränderungen für beide Krankheitsbilder berechnet.

#### 5.5.1.4 Ergebnisse

##### 5.5.1.4.1 Testeffekte

Tabelle 43 zeigt die durchschnittlichen Testzeiten der Gruppen beim Vortest 1 und 2.

*Tabelle 43: Testeffekt: Testzeiten (s) der Patienten beim Vortest 1 und 2 sowie personengruppenabhängige Zusammenhänge ( $M \pm SD$ ;  $r$ ; \*\*\* =  $p \leq 0,001$ ;  $n = 66$ )*

Gruppe (n)	Vortest 1 (s)	Vortest 2 (s)	Korrelation (r) Vortest 1, 2
HNO Gruppe 2 (36)	15,7 ± 4,3	14,7 ± 4,1	0,94***
Neurologie Gruppe 2 (30)	22,5 ± 6,7	21,3 ± 6,3	0,95***

Zwischen Vortest 1 und 2 bestehen große Zusammenhänge hinsichtlich der Zeiten in beiden Gruppen ( $p \leq 0,001$ ; HNO:  $r = 0,94$ ; Neurologie:  $r = 0,95$ ).

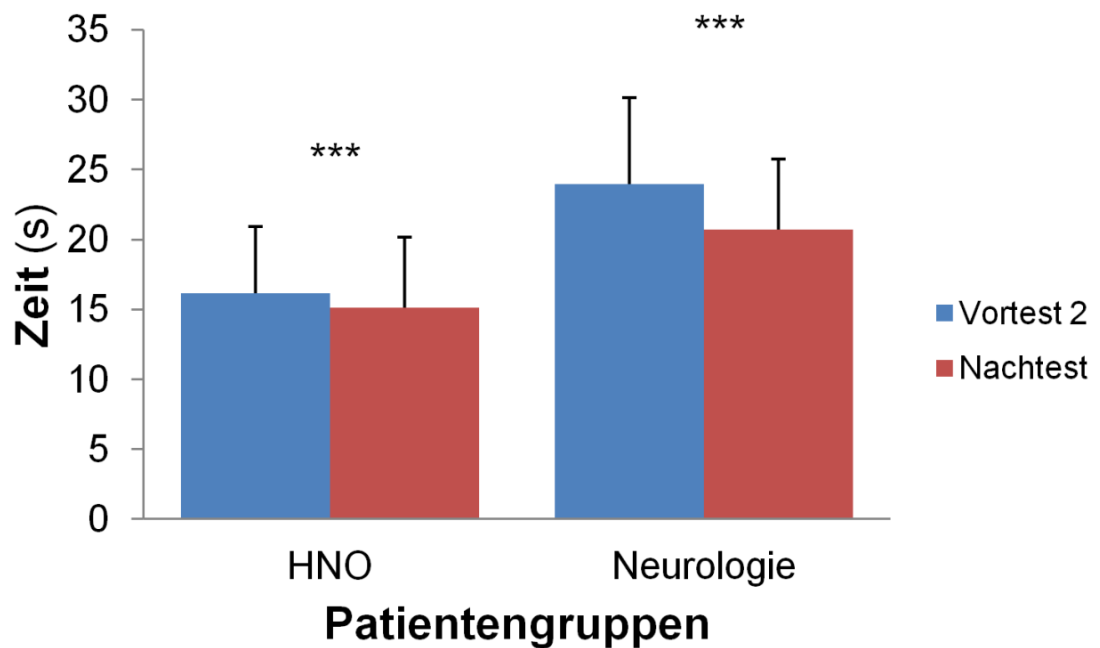
#### 5.5.1.4.2 Responsivität (nach zweiwöchiger Intervention)

In Tabelle 44 sind die durchschnittlichen Zeiten dargestellt, die die Patienten beim Vortest 2 und beim Nachtest benötigten.

*Tabelle 44: Responsivität: Testzeiten (s) der Patienten beim Vortest 2 und Nachtest ( $M \pm SD$ ;  $n = 60$ )*

Gruppe (n)	Vortest 2 (s)	Nachtest (s)
HNO (29)	16,1 $\pm$ 4,8	15,1 $\pm$ 5,0
Neurologie (31)	23,9 $\pm$ 6,2	20,7 $\pm$ 5,1

Beide Patientengruppen waren nach zweiwöchiger Intervention schneller als beim Vortest ( $p \leq 0,001$ ; Abb. 39). Patienten der HNO verbesserten sich im Mittel um rund 7 % ( $d = 0,20$ ) während sich neurologische Patienten um 15 % ( $d = 0,56$ ) verbesserten. Neun von 29 Personen mit HNO und fünf von 31 Personen mit neurologischen Erkrankungen waren langsamer als vor der Intervention.



*Abbildung 39: Unterschiede der Testzeiten (s) bei den Patientengruppen nach zweiwöchiger Intervention ( $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ ;  $n = 60$ )*



## 5.6 Validität

### 5.6.1 Kriteriumsvalidität

#### 5.6.1.1 Fragestellungen

Die Kriteriumsvalidität wird „durch den Bezug der Testpunktwerte zu Kriterienpunktwerten definiert“ (Lienert, & Raatz, 1998, S. 220). Analog zum TUG, dessen Testwerte mit den Ergebnissen externer Kriterien verglichen wurden (s. Tab. 10, S. 53), sollte überprüft werden, ob der TUG<sub>plus</sub> ebenfalls einen Bezug zu anderen Testverfahren aufweist<sup>8</sup>.

#### 5.6.1.2 Methodik

##### 5.6.1.2.1 Personenstichprobe

Die Kriteriumsvalidität wurde anhand von fünf Messverfahren (Kriterien) beurteilt. Die Stichprobengrößen unterschieden sich hierbei jeweils (Tab. 45).

*Tabelle 45: Anthropometrische Daten der Probanden, differenziert nach Validitätskriterien ( $M \pm SD$ )*

Kriterium (n)	Personengruppe (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
<b>Fahrradergometrie</b> (37)	Gesunde (37)	57 ± 14	173 ± 8	72 ± 15
<b>Timed Up and Go Test</b> (520)	Neurologie (152) HNO (180) Innere (79) Orthopädie (34) Gesunde (75)	56 ± 11	172 ± 10	79 ± 17
<b>Motorische Basisdiagnostik</b> (119)	Neurologie (67) HNO (52)	52 ± 9	174 ± 9	81 ± 17
<b>Mobilitäts-Funktionalitäts-Index</b> (226)	Neurologie (73) HNO (81) Innere (22) Orthopädie (10) Gesunde (40)	55 ± 11	172 ± 12	78 ± 18
<b>Handkraft</b> (74)	Neurologie (41) HNO (33)	52 ± 10	172 ± 15	82 ± 18

<sup>8</sup> Die Daten wurden im Rahmen einer im Arbeitsbereich durchgeführten Masterarbeit in Zusammenarbeit mit Frau Pogrzybny (2016) erhoben. Die Untersuchung und Ergebnisse wurden teilweise in den Tagungsbänden *Prävention und Rehabilitation in Zeiten der Globalisierung* (Kaczmarek et al., 2017a) und *Bewegung, Raum und Gesundheit* (Kaczmarek et al., 2017c) veröffentlicht.

#### 5.6.1.2.2 Variablenstichprobe

Die Fahrradergometrie (FE), Borg-Skala, Motorische Basisdiagnostik (MBD), Handkraft (HK), der Timed Up and Go Test (TUG) und der Mobilitäts-Funktionalitäts-Index (MFI) dienten als Außenkriterien (s. Kap. 5.1.3, S. 72).

#### 5.6.1.2.3 Ablauf der Untersuchung

Die Rekrutierung der Patienten wurde in Kapitel 5.1.2 (S. 69) beschrieben. Der Ablauf der Messung des TUG und der Handkraft wurde in Kapitel 5.1.3.3 (S. 74, 75) dargestellt.

##### *MBD*

Um lange Wartezeiten an den Teststationen bei der Fülle an Verfahren zu vermeiden und somit einen ökonomischen Ablauf zu gewährleisten, mussten die Patienten jeweils mit unterschiedlichen Stationen beginnen. Eine Gruppe begann immer mit der MBD und die andere Gruppe mit dem TUG<sub>plus</sub>. Anschließend wurden die Stationen getauscht. Die Zuteilung der Patienten zu den Gruppen sowie die Gruppenzuteilung zu den Testverfahren erfolgte randomisiert.

##### *MFI*

Die Probanden beantworteten eigenständig die Fragen zum Mobilitätsstatus (s. Kap. 5.1.3.1, S. 72).

### 5.6.1.3 Hypothesen

Hypothese 13: Die Validitätskriterien (TUG, MBD, FE, MFI, Handkraft) korrelieren untereinander ( $r > 0,40$  bzw.  $r > -0,40$ ).

Hypothese 14: Je schneller der TUG<sub>plus</sub> durchlaufen wird, 1) desto höher ist die Leistungsfähigkeit (W/kg KG), 2) desto kürzer sind die Testzeiten beim TUG, 3) desto höher die MBD- und MFI-Punktwerte und Handkraft ( $r > 0,40$  bzw.  $r > -0,40$ ).

### 5.6.1.4 Statistik

Die Zusammenhänge der Validitätskriterien (TUG, MBD, MFI, Handkraft, FE) wurden aufgrund normalverteilter Daten mittels Produkt-Moment-Korrelation nach Pearson berechnet.

Die Zusammenhänge zwischen dem TUG<sub>plus</sub> und den Kriterien wurden aufgrund nicht normalverteilter Testzeiten mittels Rangkorrelation nach Spearman berechnet.

### 5.6.1.5 Ergebnisse

#### 5.6.1.5.1 Außenkriterien – Interkorrelationen

Die Ergebnisse bei den Außenkriterien sind in Tabelle 46 dargestellt. Bei FE waren alle Teilnehmer ausbelastet (s. Kap. 5.1.3.2, S. 73).

*Tabelle 46: Ergebnisse der Validitätskriterien ( $M \pm SD$ ; TUG:  $n = 520$ ; MBD:  $n = 119$ ; MFI:  $n = 226$ ; HK:  $n = 74$ ; FE:  $n = 37$ )*

<b>TUG</b> (s)	<b>MBD</b> (Summe)	<b>MFI</b> (Summe)	<b>HK rechts</b> (N)	<b>HK links</b> (N)	<b>FE</b> (W/kg KG)
6,1 ± 1,4	12,7 ± 4,0	64,1 ± 9,0	375,7 ± 126,5	357,0 ± 132,4	2,4 ± 0,7

Bis auf die Kriterien TUG und Handkraft links, MBD und Handkraft rechts und links zeigten die Kriterien mittlere Korrelationen zueinander (Tab. 47).

*Tabelle 47: Interkorrelationen der Validitätskriterien Timed Up and Go Test (TUG:  $n = 520$ ), Motorische Basisdiagnostik (MBD:  $n = 119$ ), Mobilitäts-Funktionalitäts-Index (MFI:  $n = 226$ ), Handkraft (HK rechts; links:  $n = 74$ ) und Fahrradergometrie (FE:  $n = 37$ ;  $r$ ; \* =  $p \leq 0,05$ ; \*\* =  $p \leq 0,01$ ; \*\*\* =  $p \leq 0,001$ ; n.s. = nicht signifikant)*

	<b>TUG</b> (s)	<b>MBD</b> (Summe)	<b>MFI</b> (Summe)	<b>HK rechts</b> (N)	<b>HK links</b> (N)	<b>FE</b> (W/kg KG)
<b>TUG</b>	-	-0,60***	-0,41***	-0,23*	-0,22 <sup>n.s.</sup>	-0,49**
<b>MBD</b>	-0,60***	-	0,60***	0,24 <sup>n.s.</sup>	0,25 <sup>n.s.</sup>	-
<b>MFI</b>	-0,41***	0,60***	-	0,43***	0,39***	-
<b>HK rechts</b>	-0,23*	0,24 <sup>n.s.</sup>	0,43***	-	0,92***	-
<b>HK links</b>	-0,22 <sup>n.s.</sup>	0,25 <sup>n.s.</sup>	0,39***	0,92***	-	-
<b>FE</b>	-0,49**	-	-	-	-	-

#### 5.6.1.5.2 Zusammenhang TUG<sub>plus</sub> und Außenkriterien

Personen, die den TUG durchführten, benötigten durchschnittlich  $6,1 \pm 1,4$  s. Der TUG<sub>plus</sub> zeigt eine mittlere Korrelation von 0,64 ( $p < 0,001$ ) zum TUG (Tab. 48). Hinsichtlich der MBD-Gesamtsumme liegt ebenfalls eine mittlere Korrelation von  $r = -0,63$  ( $p \leq 0,001$ ) vor. Mittlere Zusammenhänge sind auch zur Handkraft rechts und links ( $r = -0,48$ ;  $r = -0,39$ ;  $p \leq 0,001$ ) sowie zum MFI ( $r = -0,38$ ;  $p \leq 0,001$ ) zu erkennen.

Die Probanden, die an FE teilnahmen, benötigten beim TUG<sub>plus</sub>  $14,6 \pm 4,4$  s. Bei FE erreichten männliche Probanden im Mittel eine Leistung von  $2,5 \pm 0,7$  und weibliche Probanden  $2,4 \pm 0,7$  W/kg KG.

Zwischen TUG<sub>plus</sub> (Testzeit) und FE (W/kg KG) besteht ein mittlerer Zusammenhang von  $r = -0,67$  ( $p \leq 0,001$ ; Tab. 48).

*Tabelle 48: Korrelationen zwischen den Testzeiten (s) des TUG<sub>plus</sub> und den Validitätskriterien Timed Up and Go Test (TUG:  $n = 520$ ), Motorische Basisdiagnostik (MBD:  $n = 119$ ), Mobilitäts-Funktionalitäts-Index (MFI:  $n = 226$ ), Handkraft (HK rechts, links:  $n = 74$ ) und Fahrradergometrie (FE:  $n = 37$ ;  $r$ ; \*\*\* =  $p \leq 0,001$ )*

	<b>TUG</b> (s)	<b>MBD</b> (Summe)	<b>MFI</b> (Summe)	<b>HK rechts</b> (N)	<b>HK links</b> (N)	<b>FE</b> (W/kg KG)
<b>TUG<sub>plus</sub></b> (s)	0,64***	-0,63***	-0,38***	-0,48***	-0,39***	-0,67***

## 5.6.2 Konstruktvalidität

### 5.6.2.1 Fragestellungen

In dieser Untersuchung sollte das dem Screening zugrunde liegende inhaltliche Konstrukt mittels Faktorenanalysen untersucht werden<sup>9</sup> (Bentler, & Bonett, 1980, S. 588).

Bei der Faktorenanalyse unterscheidet man zwischen der explorativen (EFA) und konfirmatorischen Faktorenanalyse (KFA). Hinsichtlich der Evaluation eines Testverfahrens ermöglicht eine EFA die Anzahl der Faktoren (Testkonstrukte) zu bestimmen, die den Indikatoren (Testergebnisse der Items) zugrunde liegen.

Die KFA (auch Strukturgleichungsmodell) untersucht die Beziehungen zwischen Indikatoren und Faktoren. Die Zuordnung der Items zu einem Faktor wird dabei durch die Korrelation zwischen den einzelnen Items bestimmt. Items, die eine hohe Korrelation zueinander aufweisen, werden dementsprechend durch denselben Faktor abgebildet.

Aufgrund der theoretischen Einbettung des Screenings in das Mobilitätskonzept der ICF (s. Kap. 2.2.2, S. 16; Abb. 3, S. 17; Abb. 16, S. 61), wurde untersucht, ob die Items bei einer Gesamtbetrachtung („Körperlich Kranke“ und „Körperlich Gesunde“) und Einzelbetrachtungen tatsächlich dieses Konstrukt abbilden.

### 5.6.2.2 Methodik

#### 5.6.2.2.1 Personenstichprobe

Zur Überprüfung der Konstruktvalidität kam der gleiche Datensatz zur Anwendung, der zur Berechnung der internen Konsistenz verwendet wurde (s. Tab. 31, S. 109).

So nahmen ebenfalls 166 Probanden (Neurologie: n = 45; HNO: n = 59; Innere Medizin: n = 25; Orthopädie: n = 26; „Gesunde“: n = 11) an der Untersuchung teil (Tab. 49).

---

<sup>9</sup> Die Untersuchung wurde teilweise auf der dvs-Tagung *Innovation & Technologie im Sport* vorgestellt und im Abstractband veröffentlicht (Kaczmarek et al., 2017b).

*Tabelle 49: Konstruktvalidität – Anthropometrische Daten der getesteten Personengruppen ( $M \pm SD$ ;  $n = 166$ )*

Personengruppe (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Neurologie (45)	♀ (19)	51 $\pm$ 6	167 $\pm$ 7	75 $\pm$ 17
	♂ (26)	54 $\pm$ 8	177 $\pm$ 7	88 $\pm$ 16
HNO (59)	♀ (32)	50 $\pm$ 9	169 $\pm$ 7	69 $\pm$ 11
	♂ (27)	53 $\pm$ 7	180 $\pm$ 7	91 $\pm$ 10
Innere Medizin (25)	♀ (11)	60 $\pm$ 6	164 $\pm$ 4	67 $\pm$ 8
	♂ (14)	67 $\pm$ 8	173 $\pm$ 5	81 $\pm$ 10
Orthopädie (26)	♀ (16)	51 $\pm$ 16	165 $\pm$ 5	71 $\pm$ 12
	♂ (10)	51 $\pm$ 15	178 $\pm$ 7	87 $\pm$ 20
„Gesunde“ (11)	♀ (8)	61 $\pm$ 7	164 $\pm$ 7	63 $\pm$ 10
	♂ (3)	55 $\pm$ 5	185 $\pm$ 4	83 $\pm$ 7

#### 5.6.2.2.2 Variablenstichprobe und Untersuchungsablauf

Der TUG<sub>plus</sub> wurde, wie in Kapitel 4.2.2 (S. 60) beschrieben, von den Probanden standardisiert durchlaufen. Die bereits für die Berechnung der internen Konsistenz benutzten Zeiten (s) für jedes Item (s. Tab. 33, S. 112) wurden für die faktorenanalytischen Berechnungen verwendet.

#### 5.6.2.3 Hypothesen

Hypothese 15: Der TUG<sub>plus</sub>, getestet anhand der einzelnen Itemzeiten, besitzt im Gesamtmodell und bei „körperlich Gesunden“ sowie „körperlich Kranken“ eine einfaktorielle Struktur mit einer guten Modellanpassung ( $RSMEA < 0,06$ ,  $CFI > 0,95$ ,  $TLI > 0,90$ ).

#### 5.6.2.4 Statistik

Zur Berechnung der Item-Reliabilität diene Cronbachs alpha ( $\alpha$ ). Zur Testung der Konstruktvalidität des Gesamtmodells, „Körperlich Kranke“ und „Körperlich Gesunde“ wurden die Itemzeiten jeweils einer EFA und KFA unterzogen.

#### Voraussetzungen

Folgende Voraussetzungen zur Durchführung dieser Analysen wurden geprüft (s. Kap. 5.1.4, S. 78):

- Itemqualität

*Alle Modelle:* Es liegt ein angemessenes Verhältnis der Testteilnehmer und der Itemanzahl vor. Bei einer Stichprobengröße von

$n = 166$  zeigt jedes Item eine akzeptable Kommunalität (alle Items mit Kommunalitäten  $h_i^2 \geq 0,55$ ; Tab. 50).

*Tabelle 50: Kommunalitäten der Items für das Gesamtmodell ( $n = 166$ ), „Körperlich Gesunde“ ( $n = 31$ ) und „Körperlich Kranke“ ( $n = 135$ )*

Item	Kommunalitäten ( $h_i^2$ )		
	Gesamtmodell	„Körperlich Gesunde“	„Körperlich Kranke“
Aufstehen (Bank), Gehen	0,69	0,59	0,68
Hindernis überwinden	0,74	0,72	0,71
Kasten umstoßen	0,58	0,69	0,55
Hinsetzen/ Aufstehen (Boden)	0,78	0,78	0,76
Kasten tragen, Hinsetzen (Bank)	0,75	0,58	0,77

- Itemkorrelation:

Tabelle 51 zeigt die Maße zur Itemkorrelationen für das Gesamtmodell, „Körperlich Kranke“ und „Körperlich Gesunde“. Der Kaiser-Meyer-Olkin Koeffizient (KMO) und der Bartlett-Test wurden gewählt, um die Eignung der Items zu testen.

Bei allen Modellen eignen sich die Items für die Durchführung von Faktorenanalysen. Zudem kann Multikollinearität ausgeschlossen werden, da alle Toleranzen größer als 0,1 sind.

*Tabelle 51: Itemkorrelationen der Modelle Gesamt ( $n = 166$ ), „Körperlich Gesunde“ ( $n = 31$ ) und „Körperlich Kranke“ ( $n = 135$ ; KMO, Bartlett-Test, MSA, Multikollinearität)*

	Gesamtmodell	„Körperlich Gesunde“	„Körperlich Kranke“
<b>KMO</b>	0,88	0,80	0,86
<b>Bartlett-Test</b>	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$
<b>Multikollinearität</b>	keine (Toleranzen 0,33 - 0,58)	keine (Toleranzen 0,19 - 0,53)	keine (Toleranzen 0,36 - 0,61)

### Explorative Faktorenanalyse

Die EFA soll Informationen über die zugrundeliegende Struktur des TUG<sub>plus</sub> geben. Hierbei wurde die Hauptkomponentenanalyse durchgeführt. Die Beziehungen zwischen den Items und den Faktoren sind als Faktorladungen ( $\lambda$ ) dargestellt. Die Ausgabe von Korrelationen kleiner 0,3 wurde unterdrückt.

Das Kaiser-Guttman-Kriterium diente zur Vorhersage der Anzahl an Faktoren. Dieses Kriterium war ein akzeptabler Prädiktor, da die Anzahl der Items größer als die Anzahl der ermittelten Faktoren war (Yeomans, & Golder, 1982). Zeigt ein Faktor einen Eigenwert größer 1, so erklärt dieser mehr Varianz als eine Variable. Als zusätzliches subjektives Ausschlusskriterium wurde der Scree-Test verwendet.

### Konfirmatorische Faktorenanalyse

Die Maximum-Likelihood-Schätzung wurde verwendet. Zur Prüfung der Modellanpassung dienten der Chiquadrat-Test ( $\chi^2$ ) und Fit Indizes (s. Kap. 5.1.4, S. 78; Tab. 19, S. 78). Für die verwendete Stichprobengröße charakterisieren RSMEA-Werte kleiner 0,06, CFI-Werte größer 0,95, und TLI-Werte größer 0,95 eine gute Modellanpassung (Hu, & Bentler, 1999).

#### 5.6.2.5 Ergebnisse

In Tabelle 52 sind die durchschnittlich benötigten Zeiten für die Items dargestellt.

*Tabelle 52: Benötigte Testzeiten (s) für die Absolvierung der jeweiligen Items beim Gesamtmodell (n = 166), bei „Körperlich Gesunden“ (n = 31) und „Körperlich Kranken“ (n = 135; s; M  $\pm$  SD)*

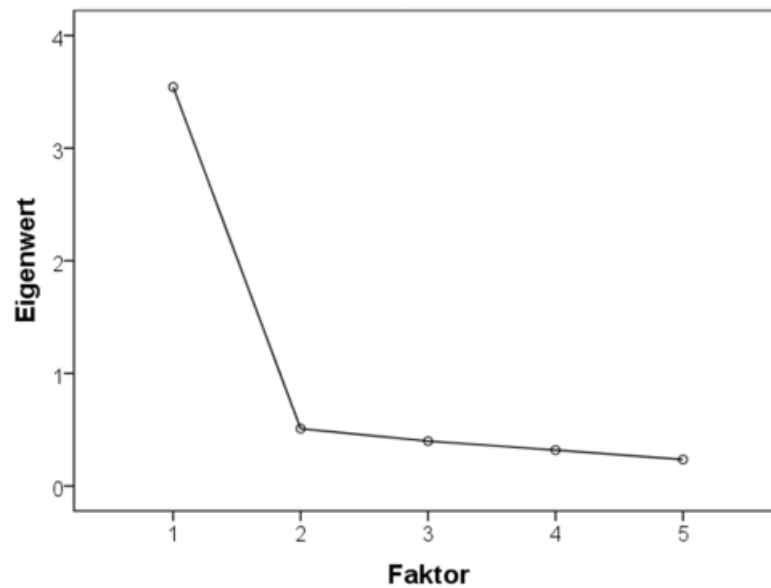
Item	Gesamtmodell	„Körperlich Gesunde“	„Körperlich Kranke“
Aufstehen (Bank), Gehen	2,5 $\pm$ 0,7	2,1 $\pm$ 0,4	2,5 $\pm$ 0,7
Hindernis überwinden	3,6 $\pm$ 1,3	2,7 $\pm$ 0,9	3,8 $\pm$ 1,3
Kasten umstoßen	3,5 $\pm$ 1,9	2,6 $\pm$ 0,7	3,8 $\pm$ 2,0
Hinsetzen/ Aufstehen (Boden)	7,0 $\pm$ 2,4	5,4 $\pm$ 1,5	7,3 $\pm$ 2,4
Kasten tragen, Hinsetzen (Bank)	8,2 $\pm$ 2,3	7,1 $\pm$ 2,1	8,4 $\pm$ 2,3



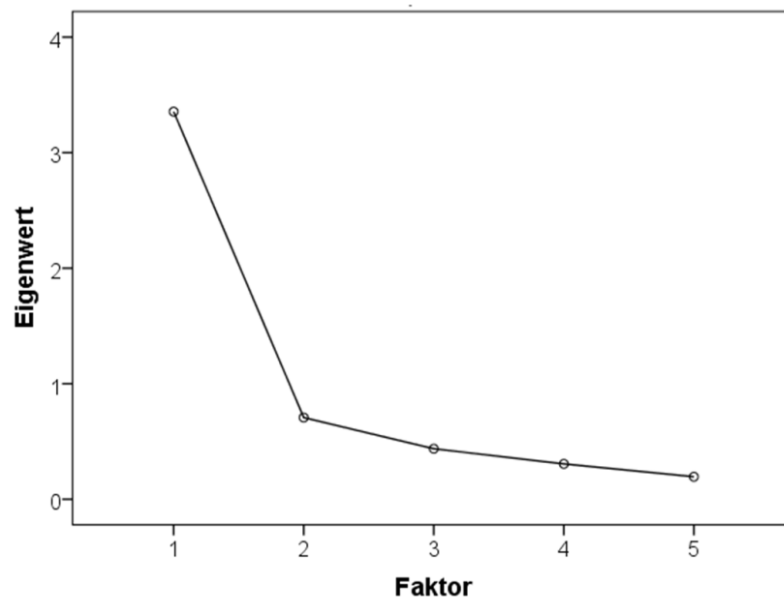
### 5.6.2.5.1 Explorative Faktorenanalyse

Alle drei Modelle zeigten hohe Reliabilitäten (Gesamtmodell:  $\alpha = 0,85$ ; „Körperlich Gesunde“:  $\alpha = 0,80$ ; „Körperlich Kranke“:  $\alpha = 0,85$ ). Das Kaiser-Guttman-Kriterium lieferte jeweils nur einen Faktor mit einem Eigenwert größer 1 (Abb. 40).

a)



b)



c)

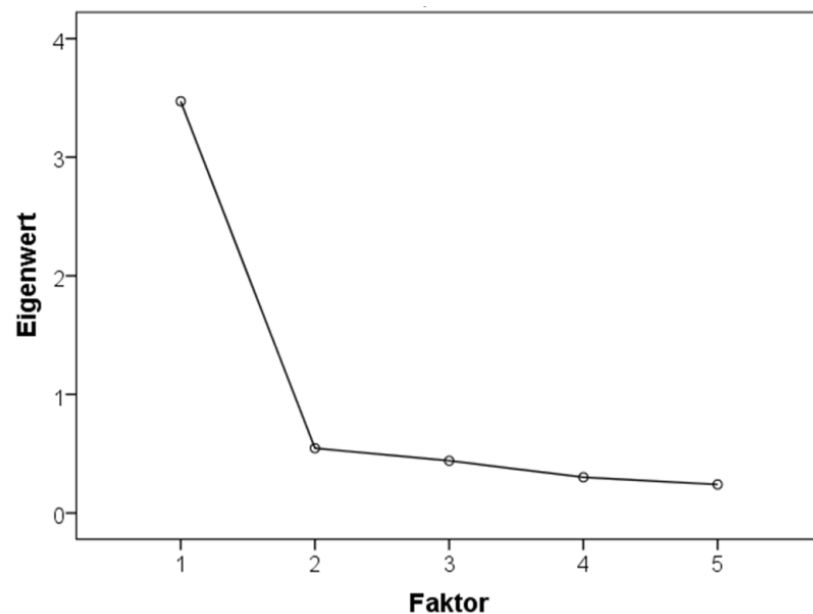


Abbildung 40: Scree-Test zur Bestimmung der Anzahl an Faktoren: a) Gesamtmodell:  $n = 166$ ; b) „Körperlich Gesunde“:  $n = 31$ ; c) „Körperlich Kranke“:  $n = 135$

- **Gesamtmodell:** Der hier gemessene Faktor erklärt 71 % der Varianz. Die Items zeigen hohe Faktorladungen ( $\lambda = 0,68 - 0,88$ ). Das Item „Kasten umstoßen“ hat die geringste und „Hinsetzen/Aufstehen (Boden)“ die größte Faktorladung.
- **„Körperlich Gesunde“:** Der hier berechnete Faktor trägt 67 % zur Varianzaufklärung bei. Die Items zeigen hohe Faktorladungen ( $\lambda = 0,76 - 0,88$ ). Im Gegensatz zum Gesamtmodell hat hier das Item „Kasten tragen, Hinsetzen (Bank)“ die geringste Faktorladung. Wie beim Gesamtmodell zeigte sich die größte Faktorladung beim Item „Hinsetzen/Aufstehen (Boden)“
- **„Körperlich Kranke“:** Der Faktor erklärt 69 % der Varianz. Die Items zeigen auch hier hohe Faktorladungen ( $\lambda = 0,74 - 0,88$ ). Wie beim Gesamtmodell läßt das Item „Kasten umstoßen“ am niedrigsten auf dem Faktor. Die höchste Faktorladung zeigt sich beim Item „Kasten tragen, Hinsetzen (Bank)“.

Die Faktorladungen der Items sind im Folgenden differenziert nach den Modellen dargestellt (Tab. 53).

*Tabelle 53: Faktorladungen ( $\lambda$ ) für das Gesamtmodell ( $n = 166$ ), „Körperlich Gesunde“ ( $n = 31$ ) und „Körperlich Kranke“ ( $n = 135$ )*

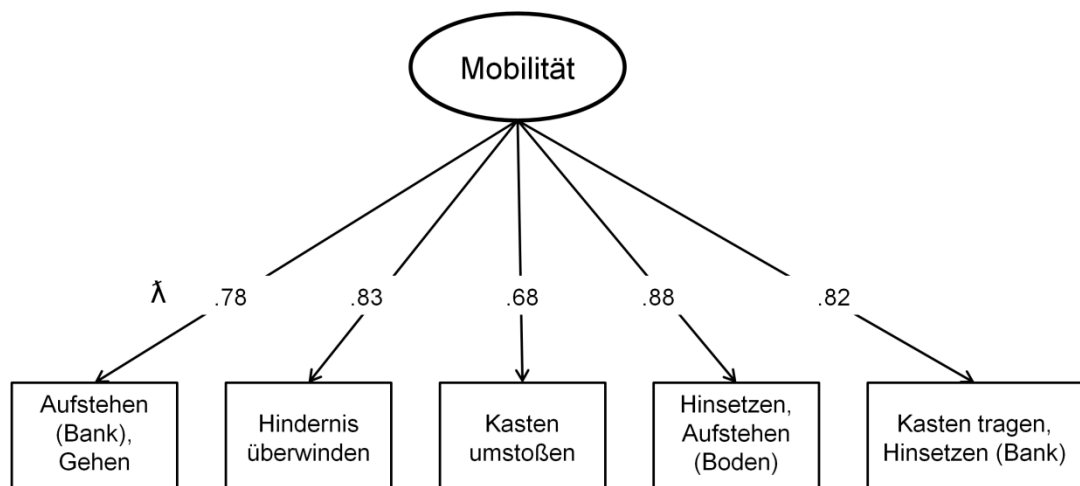
Item	Gesamtmodell	„Körperlich Gesunde“	„Körperlich Kranke“
Aufstehen (Bank), Gehen	0,78	0,77	0,82
Hindernis überwinden	0,83	0,85	0,85
Kasten umstoßen	0,68	0,83	0,74
Hinsetzen/ Aufstehen (Boden)	0,88	0,88	0,87
Kasten tragen, Hinsetzen (Bank)	0,82	0,76	0,88

Zur weiteren konfirmatorischen Betrachtung wird aufgrund der gefundenen einfaktoriellen Struktur in allen drei Modellen der Faktor als „Mobilität“ bezeichnet.

### 5.6.2.5.2 Konfirmatorische Faktorenanalyse

Der  $\chi^2$ -Wert ist in allen Modellen nicht signifikant (Abb. 41). Weiterhin ergaben sich gute Anpassungen anhand der Fit Indizes für das Gesamtmodell und „Körperlich Kranke“. Für „Körperlich Gesunde“ haben die Indizes schlechtere Anpassungen.

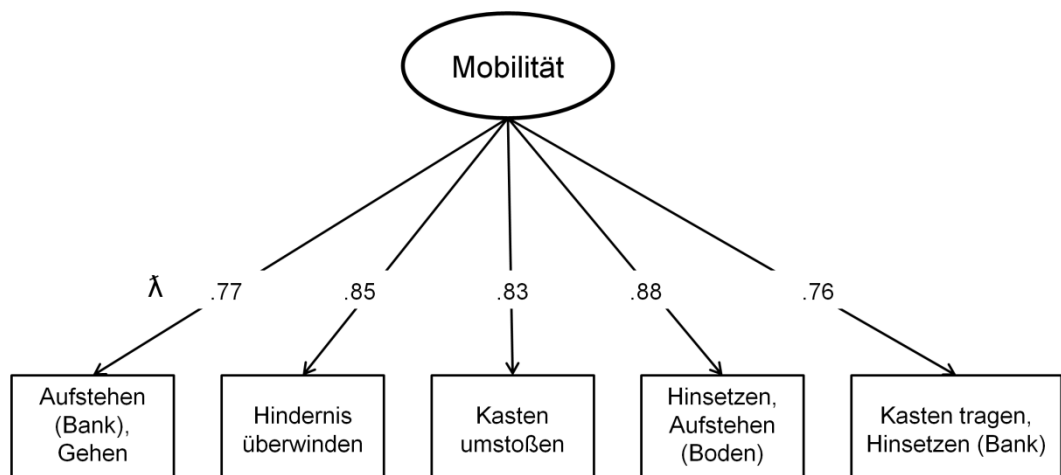
#### a) Gesamtmodell



**Modelltest:**  $\chi^2 = 6,9$ ; df = 5; p = 0,23

**Fit Indizes:** CFI = 0,99; TLI = 0,99; RMSEA = 0,05

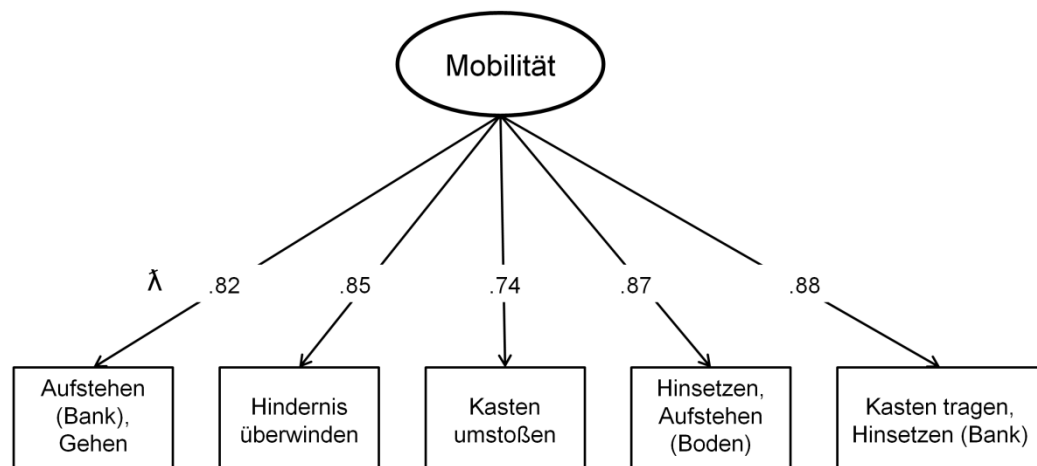
#### b) „Körperlich Gesunde“



**Modelltest:**  $\chi^2 = 10,7$ ; df = 5; p = 0,06

**Fit Indizes:** CFI = 0,92; TLI = 0,76; RMSEA = 0,04

## c) „Körperlich Kranke“



**Modelltest:**  $\chi^2 = 10,3$ ;  $df = 5$ ;  $p = 0,07$

**Fit Indizes:** CFI = 0,99; TLI = 0,96; RMSEA = 0,04

*Abbildung 41: Ergebnisse der konfirmatorischen Analyse ( $\lambda$ , Modelltest und Fit Indizes): a) Gesamtmodell:  $n = 166$ ; b) „Körperlich Gesunde“:  $n = 31$ ; c) „Körperlich Kranke“:  $n = 135$*

## 5.7 Grenzwertwertbestimmung

### 5.7.1 Ergebnisverteilung

#### 5.7.1.1 Fragestellungen

Ausgehend von der in Kapitel 3.3.4 (S. 41) beschriebenen Grenzwertproblematik, sollte im Folgenden ein Grenzwert für die Zeiten des TUG<sub>plus</sub> bestimmt werden. So galt es zu überprüfen, ob anhand von Majoritäts- bzw. Minimalnormen (s. Kap. 3.3.4, S. 41, 42) ein Grenzwertbereich definiert werden kann.

#### 5.7.1.2 Methodik

##### 5.7.1.2.1 Personenstichprobe

Nur Probanden der Gesamtstichprobe, die alle Items absolvierten (s. Tab. 28, S. 97), wurden in die Berechnungen eingeschlossen (Tab. 54).

*Tabelle 54: Ergebnisverteilung – Anthropometrische Daten der Gesamtstichprobe, differenziert nach Personengruppen ( $M \pm SD$ ;  $n = 670$ )*

Personengruppen (n)	Geschlecht	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Neurologie (152)	♀ 59	52 ± 9	167 ± 6	73 ± 15
	♂ 93	55 ± 10	176 ± 12	88 ± 16
HNO (181)	♀ 87	50 ± 10	166 ± 6	71 ± 13
	♂ 97	53 ± 10	179 ± 7	89 ± 13
Innere (109)	♀ 58	64 ± 8	163 ± 5	66 ± 11
	♂ 51	66 ± 7	175 ± 7	84 ± 13
Orthopädie (69)	♀ 39	58 ± 13	165 ± 7	68 ± 12
	♂ 30	58 ± 12	177 ± 5	85 ± 15
„Gesunde“ (148)	♀ 92	57 ± 11	165 ± 7	65 ± 15
	♂ 56	60 ± 12	180 ± 6	83 ± 14
Sportstudierende (11)	♀ 4	22 ± 3	167 ± 7	58 ± 8
	♂ 7	23 ± 3	179 ± 6	77 ± 11

##### 5.7.1.2.2 Variablenstichprobe und Ablauf der Untersuchung

Zur Bestimmung eines Grenzwertes durchliefen alle Probanden nach einem Probedurchlauf den TUG<sub>plus</sub> (s. Kap. 4.2.2, S. 60).

##### 5.7.1.3 Hypothesen

Über die Fragestellung (s. o.) hinaus gehende Hypothesen wurden nicht formuliert, da die Bestimmung von Majoritäts- bzw. Minimalnormen lediglich deskriptiv erfolgt.

#### 5.7.1.4 Statistik

Zur Bestimmung möglicher Grenzwerte wurde aufgrund nicht normalverteilter Zeiten ein Boxplot erstellt. Hierzu wurden Median, Minimum- und Maximumwerte sowie unteres und oberes Quartil berechnet.

#### 5.7.1.5 Ergebnisse

Der Median der Testzeiten der Gesamtstichprobe liegt bei einer Zeit von 20,4 s. 25 % der Probanden benötigten weniger als oder genau 16,0 s (unteres Quartil). Weitere 25 % der Personen benötigten weniger als oder genau 26,5 s (oberes Quartil; Abb. 42).

Der Interquartilsabstand liegt somit zwischen 16,0 s und 26,5 s.

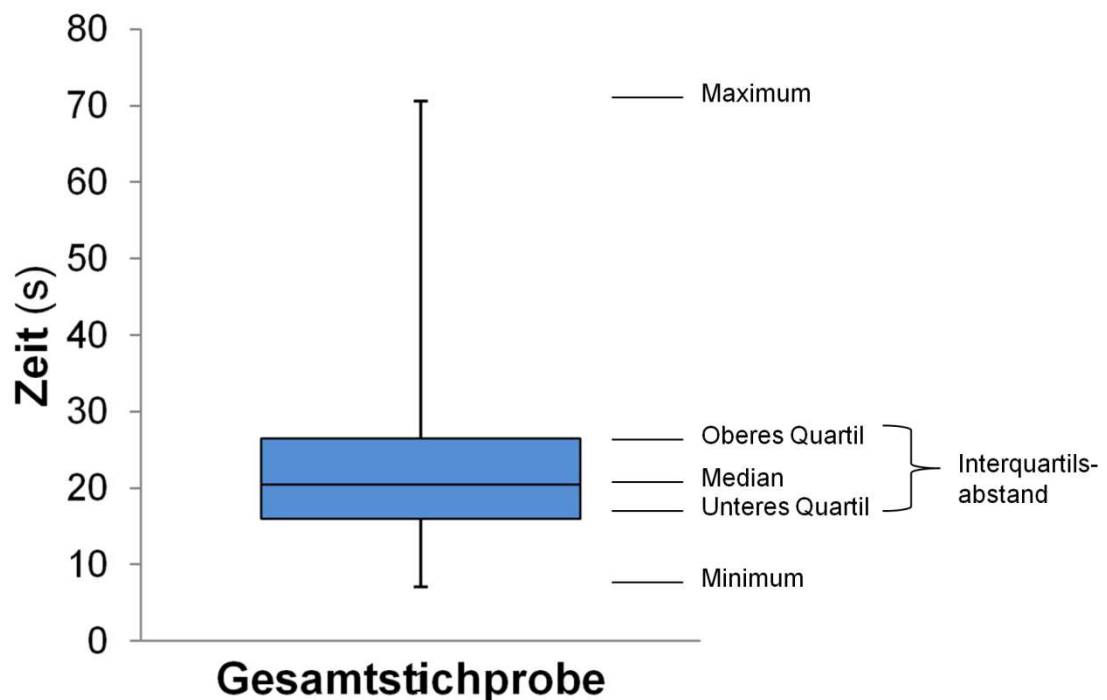


Abbildung 42: Verteilung der Testzeiten (s) der Gesamtstichprobe ( $n = 670$ )

### 5.7.2 Diagnostische Güte

Hinsichtlich der diagnostischen Güte sollten unterschiedliche Referenzstandards getestet werden, um Grenzwerte, die in der internationalen Literatur auch als *Cut-Off-Werte* bezeichnet werden, zu spezifischen Fragestellungen zu finden (s. Kap. 3.3.3, S. 37). Die Güte des Screenings wurde anhand der Referenzstandards „motorische Leistungsfähigkeit“ und „Gesundheitszustand“ überprüft. Die beiden Standards wurden zunächst separat betrachtet, anschließend zusammengeführt und gemeinsam ausgewertet sowie jeweils Cut-Off-Werte bestimmt (Abb. 43). Durch Zusammenfügung beider Standards sollten die körperlichen Beeinträchtigungen als Bedingung verschärft werden. So galt es zu überprüfen, ob der TUG<sub>plus</sub> auch bei kombinierten Einschränkungen trennscharf ist.

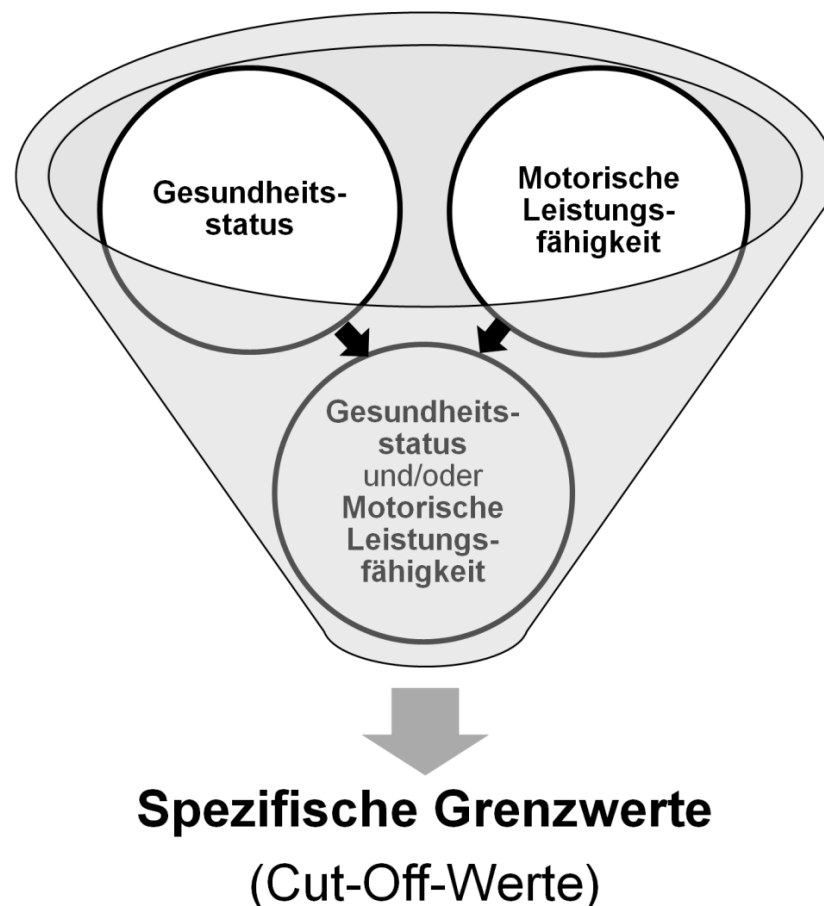


Abbildung 43: Bestimmung von spezifischen Grenzwerten anhand verschiedener Referenzstandards – Gesundheitsstatus und/oder Motorische Leistungsfähigkeit



### 5.7.2.1 Fragestellungen

*Referenzstandard „Gesundheitsstatus“:* Zunächst sollte überprüft werden, ob Personen ohne (Gesunde, HNO-Patienten) von Personen mit die Motorik beeinflussenden Beeinträchtigungen (Neurologie-, Innere-, Orthopädiepatienten) anhand des TUG<sub>plus</sub> mittels eines Cut-Off-Wertes unterschieden werden können<sup>10</sup>. Weiterhin sollte die Frage beantwortet werden, ob diejenigen, die als „Richtig-Positive“ klassifiziert werden, auch tatsächlich Probleme hinsichtlich Mobilität und motorischer Fähigkeiten aufweisen.

*Referenzstandard „Motorische Leistungsfähigkeit“:* Hierzu sollte untersucht werden, ob ein Zusammenhang zwischen dem TUG<sub>plus</sub> und der Leistungsfähigkeit (Bewältigen von 7 km/h bei einer stufenförmigen Belastung) besteht und ob Personen anhand der Leistungsfähigkeit beim TUG<sub>plus</sub> mittels Cut-Off-Wert identifizierbar sind.

*Gemeinsamer Referenzstandard „Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“:* Anschließend wurde untersucht, ob gesunde Personen ohne körperliche Beeinträchtigungen (Gesunde, HNO-Patienten), die bei einer stufenförmigen Belastung eine Geschwindigkeit von 7 km/h schaffen von Personen mit die Motorik beeinflussenden Beeinträchtigungen (Neurologie-, Innere-, Orthopädiepatienten), die das Tempo von 7 km/h nicht schaffen, unterschieden werden können.

### 5.7.2.2 Methodik

#### 5.7.2.2.1 Personenstichprobe

*Referenzstandard „Gesundheitsstatus“*

Von 221 Personen ohne und 363 Personen mit Beeinträchtigungen wurden Daten analysiert (Tab. 55). Das Fehlen und Vorhandensein von Beeinträchtigungen wurde hier als Referenzstandard benutzt. Der Gesundheitsstatus „körperlich gesund“ schloss gesunde Personen und HNO-Patienten ohne körperliche Beeinträchtigungen und ohne Schwindel ein. Personen mit körperlichen Beeinträchtigungen (Neurologie-, Innere-, Orthopädiepatienten) waren entweder stationär in saarländischen

---

<sup>10</sup> Die Untersuchung wurde teilweise auf dem 22. Sportwissenschaftlichen Hochschultag der dvs vorgestellt und im Abstractband veröffentlicht (Kaczmarek et al., 2017b).

Rehabilitationskliniken oder nahmen an ambulanten Rehabilitations- oder Präventivsportgruppen teil.

*Tabelle 55: Gesundheitsstatus – Anthropometrische Daten der getesteten Probanden ( $M \pm SD$ ;  $n = 584$ )*

Personengruppe (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Körperlich Kranke (Personen mit körperlichen bzw. die Motorik beeinflussenden Beeinträchtigungen; 221)	♀ (110)	55 $\pm$ 11	165 $\pm$ 6	66 $\pm$ 16
	♂ (101)	54 $\pm$ 9	179 $\pm$ 7	86 $\pm$ 13
Körperlich Gesunde (Personen ohne körperliche bzw. die Motorik beeinflussende Beeinträchtigungen; 363)	♀ (182)	55 $\pm$ 11	165 $\pm$ 6	70 $\pm$ 13
	♂ (181)	57 $\pm$ 12	177 $\pm$ 10	88 $\pm$ 15

#### *Referenzstandard „Motorische Leistungsfähigkeit“*

119 Personen konnten hinsichtlich des Standards getestet werden (Tab. 56).

*Tabelle 56: Motorische Leistungsfähigkeit – Anthropometrische Daten der getesteten Patientengruppen ( $M \pm SD$ ;  $n = 119$ )*

Personengruppe (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Neurologie (67)	♀ (25)	52 $\pm$ 7	167 $\pm$ 7	75 $\pm$ 17
	♂ (42)	53 $\pm$ 11	177 $\pm$ 5	87 $\pm$ 17
HNO (52)	♀ (26)	49 $\pm$ 10	166 $\pm$ 5	70 $\pm$ 12
	♂ (26)	52 $\pm$ 7	182 $\pm$ 7	90 $\pm$ 12

#### *Referenzstandard „Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“*

Von 119 Personen wurde der Gesundheitsstatus (körperlich gesund/ körperlich krank) und die Fähigkeit eine Geschwindigkeit von 7 km/h zu bewältigen, erhoben (Tab. 57). Da beide Kriterien als Referenzstandard verwendet werden sollten, wurden lediglich Personen eingeschlossen, die jeweils in beiden Fällen entweder zwei negative (keine körperlichen Beeinträchtigungen und Tempo geschafft:  $n = 16$ ; Alter  $50 \pm 9$  Jahre; BMI  $25 \pm 3$  kg/m<sup>2</sup>) oder zwei positive Ergebnisse (körperliche Beein-

trächtigkeit und Tempo nicht geschafft:  $n = 49$ ; Alter  $54 \pm 9$  Jahre; BMI  $28 \pm 5 \text{ kg/m}^2$ ) erzielten.

*Tabelle 57: Motorische Leistungsfähigkeit und Gesundheitsstatus – Anthropometrische Daten getesteten Personen ( $M \pm SD$ ;  $n = 65$ )*

Personengruppe (n)	Geschlecht (n)	Alter (Jahre)	Größe (cm)	Gewicht (kg)
Personen mit körperlichen Erkrankungen, unfähig zu laufen (49)	♀ (24)	$53 \pm 7$	$167 \pm 6$	$76 \pm 15$
	♂ (25)	$56 \pm 10$	$177 \pm 5$	$89 \pm 18$
Personen ohne körperliche Erkrankungen, fähig zu laufen (16)	♀ (6)	$53 \pm 13$	$165 \pm 7$	$65 \pm 4$
	♂ (10)	$49 \pm 6$	$184 \pm 8$	$87 \pm 14$

#### 5.7.2.2.2 Variablenstichprobe

##### *TUG<sub>plus</sub>*

Das Screening wurde bei allen drei Fragestellungen standardisiert mit jeweils einem Probedurchgang durchgeführt.

##### *Lauftest*

Die Leistungsfähigkeit der Probanden wurde anhand einer stufenförmigen Belastung (5, 6, 7 km/h, Stufendauer jeweils 2 Minuten) überprüft. Die Fähigkeit, zwei Minuten lang eine Geschwindigkeit von 7 km/h im Rahmen einer stufenförmigen Belastung (5, 6, 7 km/h) laufen zu können, diente als Kriterium zur Ermittlung eines Cut-Off-Wertes. Die Geschwindigkeit wurde in Anlehnung an Brisswalter und Mottet (1996) gewählt, die den Übergang vom Gehen zum Laufen bei einem Tempo von  $7,7 \pm 0,6 \text{ km/h}$  ermittelten.

##### *TUG*

Dieser Test wurde verwendet, um bei Personen, die als „körperlich krank“ identifiziert worden sind, Mobilitätsprobleme nachweisen zu können.

##### *Motorische Basisdiagnostik (MBD)*

Die MBD diente zur Aufdeckung von Defiziten hinsichtlich motorischer Fähigkeiten.

Die Stichprobe des Referenzstandards „Gesundheitsstatus“ absolvierte den TUG<sub>plus</sub> und die MBD (s. Kap. 5.1.3.3, S. 74), um zu überprüfen, ob

ein leistungsspezifischer Zusammenhang zwischen beiden Testverfahren besteht.

#### 5.7.2.2.3 Ablauf der Untersuchung

Nach Demonstration durch den Testleiter, absolvierten die Probanden einen Probedurchlauf beim TUG<sub>plus</sub>. Anschließend startete der Testlauf. Die Gesamtzeit (s) wurde gemessen.

Zur Generierung von Extremwerten wurden die Referenzstandards folgendermaßen dichotomisiert (s. Kap. 3.3, S. 32):

##### *„Gesundheitsstatus“*

Die Bedingung Fehlen und Vorhandensein körperlicher Beeinträchtigungen diene als Referenzstandard. Dies konnte mittels Selbstausskunft durch die Teilnehmer oder anhand eines medizinischen Berichts verifiziert werden. Somit wurde folgende Dichotomisierung vorgenommen:

- Negatives Ergebnis: Fehlen körperlicher Beeinträchtigungen
- Positives Ergebnis: Vorhandensein körperlicher Beeinträchtigungen

Bei Personen, die beim TUG<sub>plus</sub> ein positives Ergebnis erzielten, wurde zudem die Motorik beurteilt, um körperliche Beeinträchtigungen genauer beschreiben zu können. Hierbei dienten der klassische TUG und die MBD zur Identifizierung von Beeinträchtigungen hinsichtlich Mobilität und motorischer Fähigkeiten. Als Evaluationskriterien wurden beim TUG wie üblich die Testzeit (s) und bei der MBD die Scores der Items verwendet.

##### *„Motorische Leistungsfähigkeit“*

Die Probanden, die die 7 km/h schafften, wurden als leistungsfähig und diejenigen, die diese Geschwindigkeit nicht schafften als weniger leistungsfähig charakterisiert.

Dabei wurde bei den Probanden, die die 7 km/h-Stufe nicht geschafft haben nicht unterschieden, ob konditionelle oder motorische Gründe dafür verantwortlich waren. Die geringe Geschwindigkeit von 7 km/h wurde gewählt, da so konditionelle Aspekte als Abbruchgründe möglichst ausgeschlossen werden können. Es sollte lediglich die motorische Fähigkeit „Laufen können“ getestet werden.

- Negatives Ergebnis: Leistungsfähige Personen (7 km/h geschafft)
- Positives Ergebnis: Weniger leistungsfähige Personen (7 km/h nicht geschafft)

#### *„Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“*

Hier wurden die Fähigkeit und das Fehlen bzw. Vorhandensein körperlicher Beeinträchtigungen zu einem gemeinsamen Referenzstandard zusammengefasst, der wie folgt dichotomisiert wurde:

- Negatives Ergebnis: Fehlen körperlicher Beeinträchtigung und 7 km/h geschafft
- Positives Ergebnis: Vorhandensein körperlicher Beeinträchtigung und 7 km/h nicht geschafft

#### 5.7.2.3 Hypothesen

- Hypothese 16: Der TUG<sub>plus</sub> erfüllt anhand des Referenzstandards „Gesundheitsstatus“ die Kriterien für die diagnostische Güte (AUC > 0,8; LR+ 5 - 10; LR- = 0,1 - 0,2; RI > 0,6).
- Hypothese 17: Personen, die beim TUG<sub>plus</sub> langsamer als der berechnete Cut-Off-Wert beim Referenzstandard „Gesundheitsstatus“ sind, zeigen tatsächlich Mobilitäts- und Motorikprobleme mittels TUG und MBD.
- Hypothese 18: Der TUG<sub>plus</sub> erfüllt anhand des Referenzstandards „Motorische Leistungsfähigkeit“ die Kriterien für die diagnostische Güte (AUC > 0,8; LR+ 5 - 10; LR- = 0,1 - 0,2; RI > 0,6).
- Hypothese 19: Der TUG<sub>plus</sub> erfüllt anhand des Referenzstandards „Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“ die Kriterien für die diagnostische Güte (AUC > 0,8; LR+ 5 - 10; LR- = 0,1 - 0,2; RI > 0,6).

#### 5.7.2.4 Statistik

Der Vergleich der 10. und 90. Perzentile und das 95%-Konfidenzintervall (95%-CI) wurden als erste Anhaltspunkte für einen Cut-Off-Wert herangezogen.

Die diagnostische Güte wurde bestimmt durch die Receiver-Operating-Characteristic-Curve (ROC-Curve), Sensitivität, Spezifität, Positive und Negative Prädiktive Werte, Area Under The Curve (AUC), Youden-

Index, Ratz-Index (RI) und Likelihood-Ratio (LR; Bossuyt et al., 2003; Zhou et al., 2002).

Die Überprüfung auf Unterschied der Zeiten hinsichtlich der Geschwindigkeit erfolgte mittels t-Test (Normalverteilung). Aufgrund nicht normalverteilter Zeiten wurde zur Überprüfung auf Unterschied zwischen den Testzeiten ( $TUG_{plus}$ ) von Personen mit und ohne körperliche Beeinträchtigungen der Mann-Whitney-U-Test verwendet. Die Überprüfung auf Unterschied des Modells „Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“ erfolgte mittels t-Test (normalverteilte Zeiten).

Aufgrund normalverteilter Testzeiten (TUG) wurden mögliche Unterschiede hinsichtlich der Mobilität durch einen t-Test berechnet.

### 5.7.2.5 Ergebnisse

#### 5.7.2.5.1 Referenzstandard „Gesundheitsstatus“

Die als körperlich gesund definierte Gruppe (Gesunde und HNO-Patienten; s. Tab. 55, S. 140;  $18,3 \pm 6,1$  s) benötigte durchschnittlich weniger Zeit beim  $TUG_{plus}$  ( $p \leq 0,001$ ;  $d = 0,98$ ) als Personen mit körperlichen Beeinträchtigungen (Neurologie-, Innere-, Orthopädiepatienten;  $25,4 \pm 8,3$  s; Abb. 44).

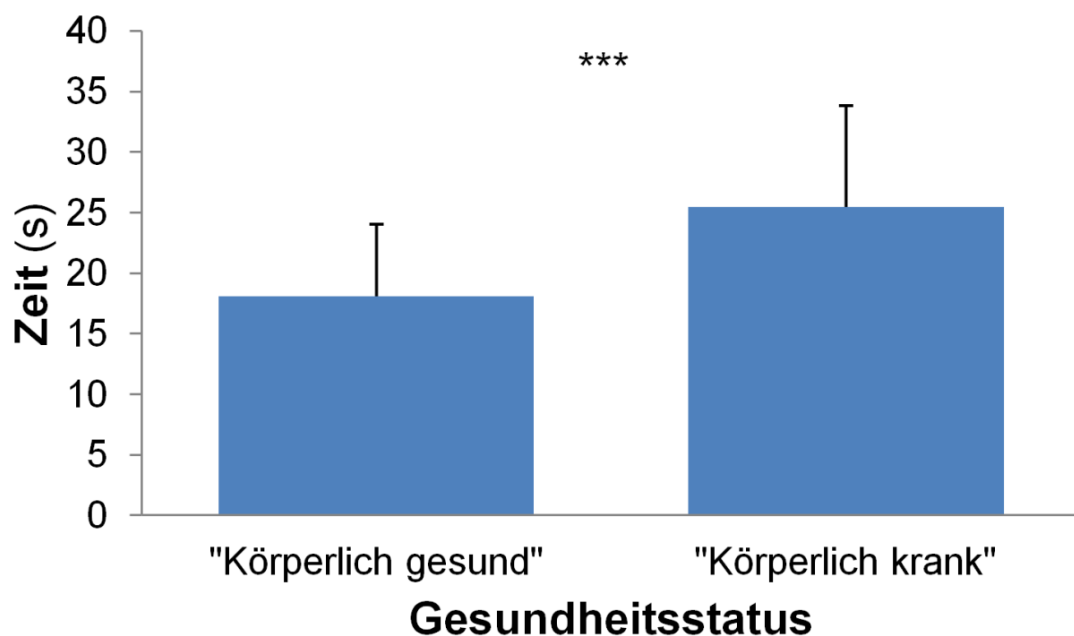


Abbildung 44: Testzeiten (s) des  $TUG_{plus}$  von Personen mit und ohne körperlichen Erkrankungen ( $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ ;  $n = 584$ )

### Cut-Off-Wert

Die 10. bis 90. Perzentile für Zeiten des Screenings von Personen ohne körperliche Erkrankungen waren 11,4 bis 27,1 s (95%-CI: 17,4 s; 19,1 s) und 16,2 bis 37,2 s (95%-CI: 24,7 s; 26,4 s) für Personen mit körperlichen Beeinträchtigungen.

Tabelle 58 zeigt die diagnostischen Gütekriterien bezüglich einer Differenzierung zwischen Personen mit und ohne körperliche Beeinträchtigungen bei Cut-Off-Werten zwischen 15 und 30 Sekunden.

*Tabelle 58: Diagnostische Güte – Differenzierung von Personen mit und ohne körperliche Beeinträchtigungen bei Cut-Off-Werten zwischen 15 und 30 Sekunden (Sensitivität – SEN, Spezifität – SPE, Positiver/Negativer Prädiktiver Wert – PPV/NPV, Area Under the Curve – AUC, Youden-Index – YI, Ratz-Index – RI; n = 584)*

TUG <sub>plus</sub> (s)	SEN	SPE	PPV	NPV	AUC	YI	RI
15	0,95	0,32	0,70	0,81	0,37	0,27	0,69
16	0,94	0,38	0,71	0,79	0,41	0,32	0,66
17	0,89	0,42	0,72	0,71	0,46	0,31	0,53
18	0,85	0,48	0,73	0,67	0,55	0,33	0,46
19	0,80	0,59	0,76	0,65	0,59	0,39	0,43
20	0,76	0,64	0,78	0,62	0,65	0,40	0,41
21	0,69	0,72	0,80	0,59	0,67	0,41	0,47
22	0,64	0,75	0,81	0,56	0,69	0,39	0,50
23	0,59	0,78	0,81	0,54	0,70	0,37	0,51
24	0,53	0,81	0,82	0,51	0,72	0,34	0,53
25	0,47	0,85	0,83	0,49	0,73	0,32	0,56
26	0,41	0,87	0,84	0,47	0,74	0,28	0,58
27	0,37	0,90	0,85	0,46	0,75	0,27	0,61
28	0,34	0,91	0,86	0,46	0,76	0,25	0,63
29	0,30	0,94	0,89	0,45	0,76	0,24	0,70
30	0,24	0,95	0,89	0,43	0,76	0,19	0,71

Wird der höchste Youden-Index (0,41) als Kriterium für die diagnostische Güte gewählt, resultiert eine TUG<sub>plus</sub>-Zeit von 21 s (95%-CI: Personen ohne körperliche Beeinträchtigungen: 17 bis 19 s; Personen mit körperlichen Beeinträchtigungen: 25 bis 26 s).

28 % der Personen ohne und 69 % der Personen mit körperlichen Beeinträchtigungen benötigten mehr als 21 s zur Absolvierung aller TUG<sub>plus</sub>-Items.

69 % der Personen, die eine Testzeit (TUG<sub>plus</sub>) langsamer als 21 s aufwiesen, haben tatsächlich körperliche Beeinträchtigungen (Richtig-Positive). Bei diesem Cut-Off-Wert werden zudem Personen ohne kör-

perliche Beeinträchtigungen auch als solche identifiziert (Richtig-Negative).

Zusätzlich wurde der Ratz-Index verwendet, um zu überprüfen, ob die Trefferquote des Screenings die Zufallstrefferquote übersteigt. Bei einem Cut-Off-Wert von 21 s führt der TUG<sub>plus</sub> zu einer leichten Verbesserung im Vergleich zur Zufallstrefferquote (Ratz-Index = 0,47; Loeber, & Dishion, 1983, s. Tab. 58, S. 145).

Die diagnostische Evidenz wurde durch die positive Likelihood-Ratio (LR+) und negative Likelihood-Ratio (LR-) bestimmt. Das 95%-CI der LR+ beträgt 2,6 bis 4,5 und des LR- von 0,4 bis 0,6. Bei einem Cut-Off-Wert von 21 s resultieren ein LR+ von 2,5 und ein LR- von 0,4.

#### Motorische Probleme

Es zeigt sich weiterhin, dass Personen, die den Parcours schneller als 21 s absolvierten (P<sub>-schnell</sub>) ebenfalls durchschnittlich weniger Zeit ( $p \leq 0,001$ ;  $d = 1,2$ ) für den TUG benötigen ( $n = 264$ ;  $5,3 \pm 1,0$  s) als Personen, die langsamer als 21 s waren (P<sub>-langsam</sub>;  $n = 287$ ;  $6,8 \pm 1,4$  s; Abb. 45).

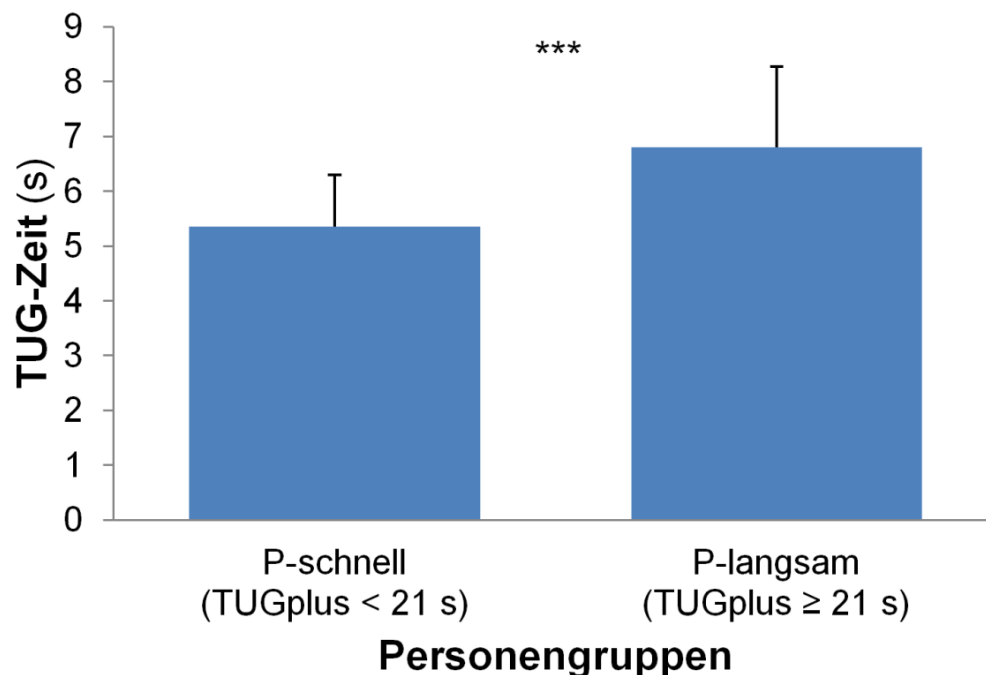


Abbildung 45: Testzeiten (s) des TUG von Personen, die weniger als 21 s (P<sub>-schnell</sub>;  $n = 264$ ) oder genau bzw. mehr als 21 s (P<sub>-langsam</sub>;  $n = 287$ ) für den TUG<sub>plus</sub> benötigen ( $M \pm SD$ )



Weiterhin konnten P<sub>-schnell</sub> (n = 55) durchschnittlich mehr MBD-Items lösen (83 %) als P<sub>-langsam</sub> (n = 64; 60 %). Der größte Unterschied zeigt sich hinsichtlich des Items „Laufen mit 7 km/h“. 78 % der P<sub>-schnell</sub> und 31 % der P<sub>-langsam</sub> lösten diese Aufgabe (Tab. 59).

*Tabelle 59: Prozentualer Anteil gelöster MBD-Items von Personen, die weniger und mehr als 21 s für den TUG<sub>plus</sub> benötigten (Referenzstandard „Gesundheitsstatus“; n = 119)*

Motorische Fähigkeit	MBD-Item	TUG <sub>plus</sub> -Zeit < 21 s	TUG <sub>plus</sub> -Zeit > 21 s
		Item gelöst (%)	Item gelöst (%)
Ausdauer	Gehen mit 5 km/h	95	88
	Gehen mit 6 km/h	93	77
	Laufen mit 6 km/h	86	48
	Laufen mit 7 km/h	78	31
Kraft	Schulterwegdrücken	96	91
	Einbeinaufstehen rechts	91	53
	Einbeinaufstehen links	87	55
	Hüftstrecken	88	66
	Rumpfaufrichten	70	42
Beweglichkeit	Rumpfflexion und -extension	95	77
	Ausschultern	80	69
	Beinstrecken	65	27
	Anfersen	94	75
Koordination	Achterkreisen	94	68
	Achterkreisen (geschlossene Augen)	48	25
	Balancieren und Ballprellen	94	78
	Ballumgreifen	65	50
	An der Wand entlang	76	67

#### 5.7.2.5.2 Referenzstandard „Motorische Leistungsfähigkeit“

Von 119 Patienten hatten 63 (67 % der HNO, 42 % der Neurologie) das Kriterium geschafft (Testzeit TUG<sub>plus</sub>: 19,5 ± 5,0 s). Probanden, die dieses Kriterium nicht erfüllten, waren langsamer ( $p \leq 0,001$ ) und benötigten im Mittel 28,5 ± 7,8 s (Abb. 46).

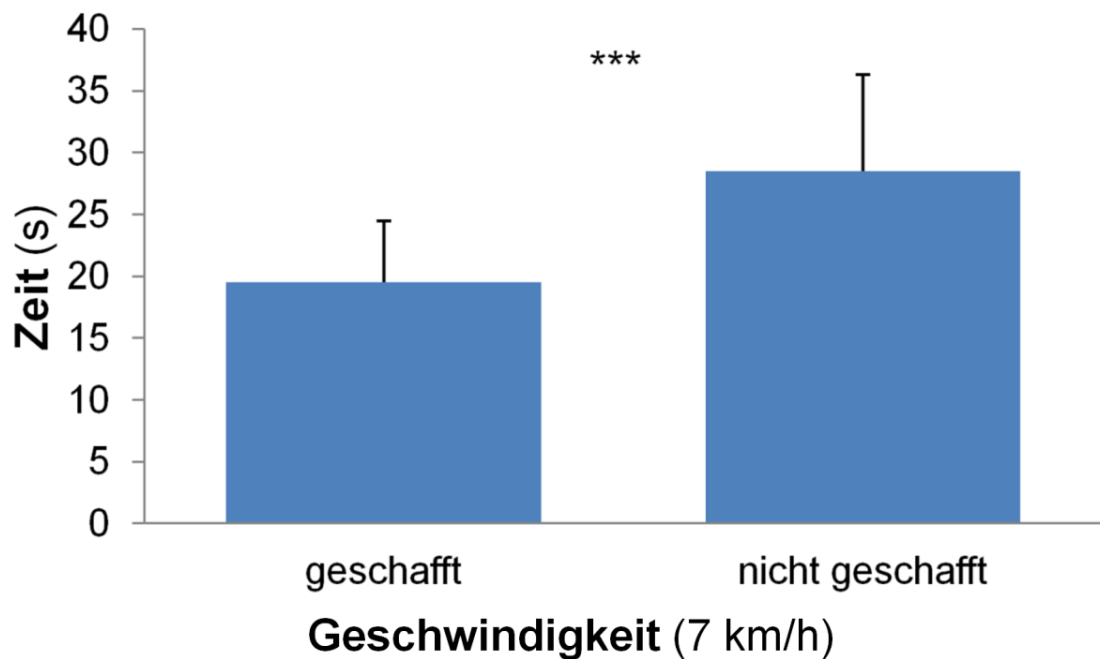


Abbildung 46: Unterschiede der Testzeiten (s) beim TUG<sub>plus</sub> bei Personen, die eine Geschwindigkeit von 7 km/h bei einer stufenförmigen Belastung schaffen ( $n = 63$ ) und solchen, die dies nicht schaffen ( $n = 56$ ;  $M \pm SD$ ; \*\*\* =  $p \leq 0,001$ )

#### Cut-Off-Wert

Der höchste Youden-Index (0,49) findet sich bei einer Zeit von 25 s (Sensitivität 68 %, Spezifität 81). Niedrigere Trennpunkte ( $< 25$  s) erhöhen die Sensitivität und den Ratz-Index, gehen jedoch zu Lasten des Youden-Index (Tab. 60). Hinsichtlich des 95%-CI liegt der Trennpunkt zwischen 22 und 28 s.

Tabelle 60: Diagnostische Güte – Differenzierung von Personen, die eine Geschwindigkeit von 7 km/h bei einer stufenförmigen Beanspruchung schaffen ( $n = 63$ ) und von solchen, die dies nicht schaffen (Testzeiten zwischen 22 und 28 s;  $n = 56$ ; Sensitivität – SEN, Spezifität – SPE, Positiver/Negativer Prädiktiver Wert – PPV/NPV, Area Under the Curve – AUC, Youden-Index – YI, Ratz-Index – RI)

TUG <sub>plus</sub> (s)	SEN	SPE	PPV	NPV	AUC	YI	RI
22	0,78	0,66	0,66	0,78	0,65	0,44	0,52
23	0,76	0,71	0,69	0,78	0,66	0,47	0,52
24	0,70	0,73	0,69	0,74	0,72	0,43	0,44
25	0,68	0,81	0,76	0,75	0,73	0,49	0,55
26	0,62	0,83	0,76	0,72	0,76	0,45	0,55
27	0,54	0,88	0,79	0,69	0,77	0,42	0,62
28	0,54	0,9	0,82	0,70	0,81	0,44	0,66

### 5.7.2.5.3 Referenzstandard „Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“

Für den Parcours benötigten Personen ohne körperliche Beeinträchtigungen, die laufen konnten  $16,0 \pm 3,8$  Sekunden (10. - 90. Perzentil: 11,2 s und 22,9 s; 95%-CI: 14,0 s, 18,0 s) und waren schneller ( $p \leq 0,001$ ) als Personen mit Beeinträchtigungen, die nicht laufen konnten  $29,1 \pm 7,7$  s (10. - 90. Perzentil: 19,5 s und 38,4 s; 95%-CI: 26,9 s, 31,3 s).

#### Cut-Off-Wert

Tabelle 61 zeigt die Gütekriterien bei Testzeiten zwischen 15 und 30 s.

*Tabelle 61: Diagnostische Güte – Differenzierung von Personen, die eine Geschwindigkeit von 7 km/h bei einer stufenförmigen Beanspruchung schaffen und körperlich nicht beeinträchtigt sind ( $n = 63$ ) von solchen, die dies nicht schaffen und gleichzeitig körperlich beeinträchtigt sind (Testzeiten zwischen 22 und 28 s;  $n = 56$ ; Sensitivität – SEN; Spezifität – SPE; Positiver/Negativer Prädiktiver Wert – PPV/NPV; Area Under the Curve – AUC; Youden-Index – YI; Ratz-Index – RI)*

TUG <sub>plus</sub> (s)	SEN	SPE	PPV	NPV	AUC	YI	RI
15	0,98	0,44	0,84	0,88	0,50	0,42	0,83
16	0,98	0,50	0,86	0,89	0,56	0,48	0,85
17	0,98	0,56	0,87	0,90	0,68	0,54	0,87
18	0,94	0,69	0,90	0,79	0,79	0,63	0,72
20	0,92	0,81	0,94	0,77	0,85	0,73	0,75
21	0,88	0,88	0,96	0,70	0,85	0,76	0,82
22	0,82	0,88	0,95	0,61	0,90	0,70	0,81
23	0,80	0,94	0,98	0,60	0,90	0,74	0,90
24	0,74	0,94	0,97	0,54	0,94	0,68	0,89
25	0,71	1	1	0,53	0,94	0,71	1
26	0,65	1	1	0,49	0,94	0,65	1
28	0,57	1	1	0,43	0,94	0,57	1
29	0,55	1	1	0,42	0,94	0,55	1
30	0,47	1	1	0,38	0,94	0,47	1

Der Youden-Index beträgt 0,76. Dies entspricht einer Testzeit von 21 s. Bei dieser Zeit erzielten 20 Personen ein negatives ( $< 21$  s) und 45 Personen ein positives Testergebnis ( $> 21$  s).

88 % der durch den Referenzstandard als körperlich unbeeinträchtigt charakterisierten Personen erzielten auch beim Screening ein negatives Ergebnis ( $< 21$  s; Spezifität – Richtig-Negativ-Rate) und 88 % der als

körperlich beeinträchtigt charakterisierten Personen erzielten ein positives Ergebnis ( $> 21$  s; Sensitivität – Richtig-Positiv-Rate).

Der Vergleich der Treffer- mit der Zufallstrefferquote ergab einen Ratz-Index von 0,82. Bezüglich der Wahrscheinlichkeitsverhältnisse resultierten ein LR+ von 7,3 und ein LR- von 0,1.

#### 5.7.2.5.4 ROC-Analysen

In Abhängigkeit des Referenzstandards ergeben sich unterschiedliche ROC-Kurvenverläufe (Abb. 47). Das Kriterium „Stufengeschwindigkeit 7 km/h“ führt bei einem Cut-Off-Wert von 25 s zu einer AUC von 0,73. Die Unterscheidung zwischen Personen ohne und mit körperlichen Beeinträchtigungen liefert eine AUC von 0,67.

Beide Standards zusammen ergeben eine AUC von 0,95 (Abb. 47).

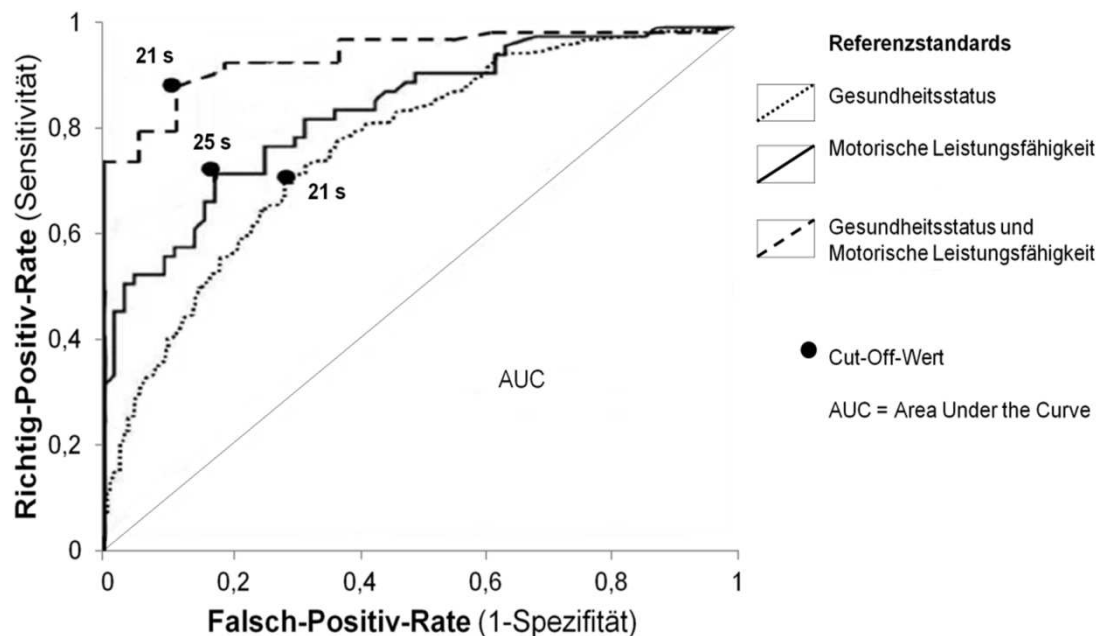


Abbildung 47: Richtig-, Falsch-Positiv-Raten und ROC-Kurven in Abhängigkeit des Referenzstandards: Berechnete Cut-Off-Werte (s) und Area Under the Curves (AUC; Cut-Off-Wert)

Die Zusammenfügung beider Referenzstandards führt also zu einer um ca. 20 % verbesserten richtigen Erkennung körperlich Beeinträchtigter (Abb. 48).

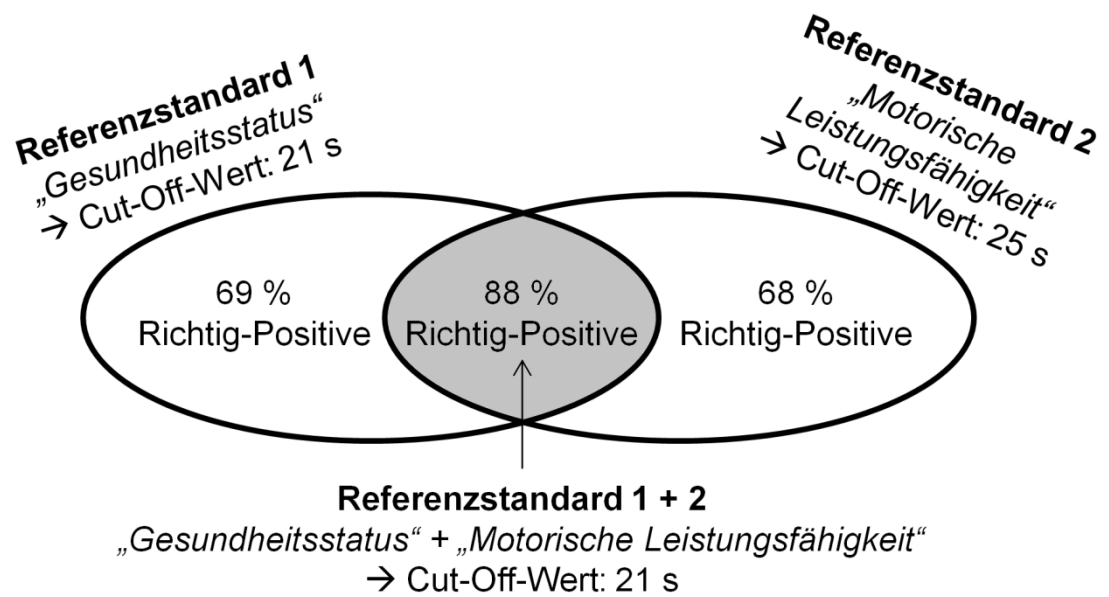


Abbildung 48: Richtig-Positiv-Raten des TUG<sub>plus</sub> in Abhängigkeit des Referenzstandards (1, 2, 1 + 2) und den ermittelten Cut-Off-Werten

## 6 Diskussion

Bei dieser Untersuchung sollte überprüft werden, inwiefern ein an Alltags-handlungen bzw. an der Internationalen Klassifikation für Funktionsfähigkeit, Behinderung und Gesundheit orientierter Timed Up and Go Test (TUG) motorisch beeinträchtigte und unbeeinträchtigte Personen unterscheidet. Dazu absolvierten 1023 Probanden mit unterschiedlichem Gesundheitsstatus den klassischen TUG sowie 764 Probanden den um zwei Items und insgesamt auf acht alltagsorientierte Bewegungsmuster erweiterten Timed Up and Go Test plus (TUG<sub>plus</sub>).

Das Interesse an validen, ökonomischen Tests im Gesundheitsbereich war die Ausgangslage zur Entwicklung eines neuen motorischen Messverfahrens (Huber, 2012; DRV, 2015b; Kaczmarek, Schwarz, & Wydra, 2017c). Neben bereits etablierten Motoriktests, die sich am klassischen Modell der motorischen Leistungsfähigkeit, der Aktivitäten des täglichen Lebens aber auch an neueren Konzepten der Kompetenzen orientieren, stellt die Konzeption ICF-basierter Verfahren eine weitere Herangehensweise dar.

Eine Testentwicklung basierend auf diesem Konzept greift die Aktualität im Gesundheitsbereich auf. So wurde in dieser Arbeit ein motorisches Screening auf der Basis der Mobilitätskategorien der ICF entwickelt (s. Abb. 3, S. 17). Diese Kategorien decken viele Aktivitätsbereiche eines Menschen ab. So kann die Erfassung der Mobilität neben der ICD zusätzliche personenbezogene Informationen liefern. Das Kodierungsschema erleichtert hierbei die Beschreibung, was eine Person motorisch leisten kann. Die Erfassung dieser sogenannten „funktionalen Gesundheit“ (s. Kap. 2.2.1, S. 16) ermöglicht jedoch keine direkten Rückschlüsse auf mögliche Verletzungen oder Erkrankungen, wie dies beispielsweise anhand von ICF-Core-Sets möglich ist (Stucki et al., 2002; s. Kap. 2.2.2, S. 18). Dort werden wichtige Aktivitätsbereiche speziell für verschiedene Erkrankungen herausgearbeitet.

Weiterhin fehlen innerhalb der ICF Hinweise zur Operationalisierung. So erschwerte die Vielfalt und Komplexität der Kodierungen, insbesondere die Gewichtung der aufgeführten Bewegungen bzw. Tätigkeiten, die Konzeption von Testaufgaben (Okochi, Utsunomiya, & Takahashi, 2005).

### Testaufgaben

In einer Voruntersuchung fand eine Selektion aus einem Itempool statt (s. Kap. 4.2.1, S. 59). Im Gegensatz zu Selektionsstrategien, die einen Ausschluss einfacher Items implizieren (Lienert, & Raatz, 1998, S. 115), wurden

solche Aufgaben mit aufgenommen, die eine Ergebnisfindung primär anhand der Testzeit und nicht anhand der Ausprägung einzelner Items erlauben.

Der klassische TUG, der hauptsächlich in der Geriatrie trennscharfe Ergebnisse liefert, diene als Testbasis (s. Kap. 4.1, S. 47) und wurde durch Hinzunahme zusätzlicher alltagsorientierter Bewegungshandlungen um zwei Items erweitert (TUG: drei; TUG<sub>plus</sub>: fünf Items mit acht Bewegungsaufgaben).

Da in Studien zum TUG die Testausstattung oft recht unterschiedlich ausfiel (Asmidawati et al., 2014; Broekmans et al., 2010; Hoogeboom et al., 2010), sollte zur besseren Standardisierung beim TUG<sub>plus</sub> auf Material zurückgegriffen werden, das in herkömmlichen Sporthallen zur Grundausstattung gehört und gut zugänglich ist (Turnbank, Matten, Turnkästen). Durch die standardisierten Größen und Gewichte ist so eine bessere Vergleichbarkeit gewährleistet.

Der TUG<sub>plus</sub> erfordert zwar einen größeren Material- (TUG: Stoppuhr, Stuhl, Pylone, TUG<sub>plus</sub>: Stoppuhr, zwei Turnbänke, drei kleine Turnkästen, drei Turnmatten) und Platzbedarf (TUG: drei Meter, TUG<sub>plus</sub>: 20 m<sup>2</sup>), beinhaltet jedoch auch mehr Bewegungsmuster mit konkretem Alltagsbezug. Die Items beinhalten Aufgaben, die inhaltlich auch den „fundamental movement milestones“ (Burton, & Miller, 1998, S. 215) zugeordnet werden können, die später von Herrmann und Gerlach (2014) als „Sich bewegen“ und „Etwas bewegen“ differenziert wurden. Zudem können die Items international vergleichbar in einem Kodierungsschema (hier Mobilitätskategorie) zugeordnet werden (WHO, 2001; s. Abb. 16, S. 61). Aus pädagogischer Sicht kann zudem durch Miteinbeziehung vielfältiger alltagsbezogener Bewegungshandlungen für die Testpersonen auch der Gegenwartsbezug erhöht werden.

Hinsichtlich der Anzahl der Durchgänge ist in Anlehnung an TUG-Studien (Asmidawati et al., 2014; Schacht, & Ringe, 2012) ein Probendurchgang nötig. Hierdurch werden die einzelnen Items erprobt, der Gesamtablauf wird deutlich und Testeffekte werden minimiert. Darüber hinaus können Personen mit höhergradigen motorischen Beeinträchtigungen (z. B. Schlaganfall-, Multiple Sklerose-, Parkinsonpatienten) hier erkannt und ggf. vom eigentlichen Test auch unter dem Aspekt der Risikominimierung ausgeschlossen werden.

Eine gute Testzeit impliziert zunächst eine ökonomische Bewältigung der einzelnen Items ohne weitere motorische Differenzierung. Durch eine zusätzliche Beurteilung der Bewegungshandlungen ist es auch bei akzeptabler Testzeit möglich motorische Defizite zu erkennen (s. Tab. 14, S. 66).

Da weiterhin die Items krankheitsunabhängig entwickelt wurden, können bestimmte, für eine Erkrankung typische Beschwerden, nicht zielsicher erfasst werden (Doyle, 2002). So ist lediglich ein erster Eindruck über die Motorik bzw. Mobilität möglich.

Es ergeben sich folgende Schlussfolgerungen:

- Der klassische TUG, der lediglich aus den Fertigkeiten „Aufstehen/Hinsetzen“, „Gehen und Richtungswechsel“ besteht, wurde durch zusätzliche motorische Alltagselemente (fünf Items mit acht motorischen Aufgaben) erweitert (s. Kap. 4.2, S. 59).
- Aufgrund eines höheren Platzbedarfs (TUG: drei Meter, TUG<sub>plus</sub>: 20 m<sup>2</sup>) ist der TUG<sub>plus</sub> im Vergleich zum TUG zwar unökonomischer, jedoch konnte der Alltagsbezug erhöht werden.
- Durch die Orientierung an der ICF kann zudem eine Einstufung anhand eines internationalen Kodierungssystems erfolgen (s. Abb. 16, S. 61).

## 6.1 Evaluation des TUG<sub>plus</sub>

### 6.1.1 Timed Up and Go Test

Podsiadlo und Richardson (1991) haben den klassischen TUG entwickelt, bei primär über 70-jährigen Probanden validiert und dabei einen Trennpunkt von 10 Sekunden für motorisch auffällige bzw. unauffällige Personen bestimmt (s. Kap. 4.1, S. 47). Bohannon (2006) stellte für den TUG im Rahmen einer Metaanalyse Normwerte für Personengruppen, beginnend bei 60-Jährigen bis 99-Jährige, auf (8 s bis 11 s; s. Tab. 8, s. 48).

Als Erweiterung zu den Befunden von Bohannon (2006) wurden in dieser Untersuchung die benötigten Testzeiten differenziert nach Personengruppe und in verschiedenen Altersgruppen gemessen.

Zunächst wurden in dieser Studie die Testzeiten in Abhängigkeit der Personen- bzw. Patientengruppe altersunabhängig auf Unterschiede überprüft (s. Abb. 20, S. 82). Es zeigte sich, dass die verschiedenen Patientengruppen und „Körperlich Gesunde“ mehr Varianz aufklären als das Alter. Mittlere bis hohe Effekte ( $d \geq 0,5$ ) ergaben sich jedoch hier lediglich beim Vergleich der jeweiligen Krankheitsbilder mit den „Körperlich Gesunden“. Die signifikant unterschiedlichen Zeiten unter den Patienten zeigten nur kleine Effekte ( $d < 0,5$ ).



In bisherigen Studien wurde insbesondere die Auswirkung neurologischer Erkrankungen auf die Testzeit beschrieben sowie die längsten Testzeiten gemessen. Schlaganfallpatienten mit höhergradigen Beeinträchtigungen benötigten zwischen 15 (Yang et al., 2006) und 28 Sekunden (Blennerhasset, & Dite, 2004; Dean et al., 2000; Salbach et al., 2004).

In der vorliegenden Untersuchung befanden sich in der Gruppe neurologischer Patienten ebenfalls Personen, die einen Schlaganfall erlitten. Im Gegensatz zu den aufgeführten Studien konnte für diese Gruppe jedoch eine deutlich geringere Testzeit von im Mittel sieben Sekunden ermittelt werden (s. Tab. 21, S. 81), was vermutlich auf den geringeren Beeinträchtigungsgrad der Erkrankung zurückzuführen ist.

Anzumerken ist, dass die getesteten Patientengruppen nicht eindeutig voneinander zu trennen sind. Aufgrund von Multimorbidität leiden die meisten Personen an mehreren Erkrankungen. Die Zuteilung zu den jeweiligen Patientengruppen erfolgte lediglich anhand des jeweiligen Hauptbeschwerdebildes.

Im Gegensatz zu den Befunden von Pondal und Ser (2008), die bei Personen im Alter zwischen 71 und 99 Jahren unterschiedliche Testzeiten bei Männern und Frauen nachweisen konnten, ergab sich bei der hier getesteten jüngeren Stichprobe keine Geschlechtsabhängigkeit.

Die getesteten Personen ( $n = 1023$ ) waren im Mittel 56 Jahre (s. Kap. 5.2.2, S. 80) und benötigten durchschnittlich 6,3 Sekunden für den Test, was mit den Befunden von Bird et al. (2012) in etwa übereinstimmt (s. Tab. 8, S. 48). Selbst die neurologischen Patienten erreichten im Mittel weniger als 7 Sekunden (s. Abb. 20, S. 82).

Bei der Gesamtstichprobe konnte trotz der Unterschiede ( $p \leq 0,01$  bzw.  $p \leq 0,001$ ) zwischen den Altersquartilen nur eine geringe Effektstärke ( $\eta^2 = 0,02$ ) berechnet werden. Die Testzeit von Personen im Alter ab 62 Jahren unterschied sich von der aller jüngerer Personengruppen ( $d < 0,5$ ), die allerdings untereinander nicht weiter differierten (s. Abb. 21, S. 84), so dass eine weitere altersspezifische Differenzierung des hier berechneten durchschnittlichen Normwertbereiches (6,2 - 6,7 s für 18 bis 87-Jährige; s. Tab. 22, S. 83) nicht notwendig erscheint.

Bei gesunden Personen wurde im Gegensatz zu den Befunden von Bohannon (2006) kein Einfluss des Alters auf die Testzeit gefunden (s. Abb. 22, S. 85).

Bei den getesteten Patientengruppen wurden lediglich bei Probanden mit HNO und inneren Erkrankungen altersabhängige Unterschiede gemessen (s. Abb. 23, S. 86; Abb. 25, S. 88). Bei diesen Krankheitsbil-

dern zeigte sich ein mit zunehmendem Alter abnehmendes Gehtempo. Bei HNO-Patienten wird dies allerdings erst ab einem Alter von 59 Jahren deutlich. Bei Patienten mit inneren Erkrankungen ergaben sich erst ab einem Alter von 64 Jahren relevant unterschiedliche Testzeiten.

Bei neurologischen und orthopädischen Patienten zeigten sich keine altersabhängigen Testzeiten (s. Abb. 24, S. 87; Abb. 26, S. 89). Entgegen den Erwartungen nahmen die Testzeiten bei orthopädischen Patienten sogar bis zu einem Alter von 65 Jahren ab (rund 6). Erst ab 66 Jahren zeigten sich erneut marginal höhere Testzeiten von rund 7 Sekunden. Bei orthopädischen Patienten scheinen somit die Erkrankungen eine höhere Auswirkung auf das Gehtempo zu haben als das Alter.

Mit durchschnittlich 6,6 Sekunden benötigten die über 62-Jährigen deutlich weniger Zeit als dies für diese Altersgruppe in der Metaanalyse von Bohannon (2006) ermittelt (60- bis 69-Jährige: 8,1 s; 70- bis 79-Jährige: 9,2 Sekunden; 80- bis 99-Jährige: 11,3 s) und in einer Untersuchung von Pondal und Ser (2008) bestätigt wurde. Im Gegensatz dazu ermittelten Tsuji et al. (2014) für 60- bis 69-Jährige mit durchschnittlich 4,9 Sekunden noch geringere Zeiten (s. Tab. 8, S. 48).

Weiterhin konnte die Feststellung von Vereeck et al. (2008), dass pro Lebensdekade die Testzeit um bis zu 0,8 Sekunden zunimmt, nicht bestätigt werden.

Zusammenfassend ist für den Altersbereich 50 bis 70 Jahren ein Trennpunkt zwischen 6 und 7 Sekunden ohne weitere Altersdifferenzierung und unabhängig vom Gesundheitsstatus anzunehmen.

Die unterschiedliche Instruktion stellt hier offenbar die größte Störquelle dar (s. Tab. 9, S. 51). So zeigte sich schon in der Erprobungsphase, dass sich die Testzeiten mit dem Hinweis, den Test „schnellstmöglich“ auszuführen, im Gegensatz zur Instruktion „schnellstmöglich und sicher“ halbiert haben.

Die Untersuchung lässt folgende Schlussfolgerungen zu:

- Für den Altersbereich von 50 bis 70 Jahren kann ein Trennpunkt zwischen 6 und 7 Sekunden ohne weitere Altersdifferenzierung und unabhängig vom Gesundheitsstatus angenommen werden.
- Innerhalb der Patientengruppen sind die neurologischen Teilnehmer langsamer als Patienten mit HNO- und inneren Erkrankungen. HNO-Patienten benötigten nicht mehr bzw. weniger Zeit als Patienten der Inneren Medizin und Orthopädie. Die Testzeiten orthopädischer und neurologischer Patienten unterschieden sich

ebenfalls nicht. **Hypothese 1** (S. 80) kann ebenfalls nur eingeschränkt angenommen werden.

- Es ergaben sich keine geschlechtsabhängigen Testzeiten. Bei den Patientengruppen HNO und Innere konnte zum Teil eine mit zunehmendem Alter geringfügig ansteigende Testzeit gemessen werden. **Hypothese 2** (S. 80) kann somit hinsichtlich des Geschlechts nicht und bezüglich des Alters nur bedingt angenommen werden.

### 6.1.2 Timed Up and Go Test plus

Vor der Überprüfung der Testgüte und diagnostischen Güte wurden zunächst die Item-Lösungshäufigkeiten untersucht. Anschließend sollten die Testzeiten in Abhängigkeit von Personengruppe, Geschlecht und Alter auf Unterschied überprüft werden.

88 % der Gesamtstichprobe konnten alle Items lösen. HNO-Patientinnen hatten im Vergleich zu den anderen Gruppen prozentual am häufigsten mindestens ein Item nicht durchgeführt und dementsprechend nur 81 % Lösungshäufigkeit erreicht (s. Tab. 24, S. 93). Bei diesen Patienten haben Begleiterscheinungen wie Schwindel zu Problemen geführt. Morris et al. (1990), Cumming et al. (2000) und Harun et al. (2015) fanden ebenfalls heraus, dass Personen mit Gleichgewichtsproblemen eine erhöhte Sturzangst haben, die bei der Ausführung alltäglicher Bewegungen hemmend wirkt. Insgesamt zeigt sich, dass insbesondere Patienten mit schwereren Erkrankungen bzw. höhergradigen Beeinträchtigungen nicht in der Lage waren alle Items des TUG<sub>plus</sub> durchzuführen (s. Kap. 5.3.5.1, Tab. 24, S. 93).

Zusammenfassend resultiert aus den Befunden, dass der TUG<sub>plus</sub> in Abhängigkeit der Personen- bzw. Patientengruppen und des Alters (s. Tab. 25, S. 94) eine geringe Gesamtschwierigkeit aufweist (Mittelwert 10 Punkte; s. Tab. 27, S. 95), was die Voraussetzung zur Verwendung des Messparameters Zeit begründet. So wurde auch das schwerste Item (Kasten umstoßen) immerhin noch von 86 % der Probanden gelöst (s. Tab. 26, S. 95).

Der Normwertbereich für Personen im Alter von 20 bis 85 Jahren liegt zwischen 18,4 und 26,2 Sekunden (exklusiv Sportstudierende). Betrachtet man geschlechtsunspezifisch die HNO-Patienten, die alle Items absolviert haben, erreichten sie ähnliche Testzeiten wie die körperlich gesunde Stichprobe (s. Abb. 28, S. 98). Möglicherweise ist dies darauf zurück zu führen, dass die Mehrzahl der HNO-Patienten, die alle Items

lösten, lediglich an Tinnitus litt, woraus kaum körperliche Einschränkungen resultieren.

Beim klassischen TUG zeigt sich anhand verschiedener Befunde ein ähnliches Bild. HNO-Patienten (Vereeck et al., 2008) sind auch bei diesem Testverfahren schneller als Patienten mit orthopädischen (Vance et al., 2012), neurologischen (Broekmans et al., 2010; Morris et al., 2001) und inneren Erkrankungen (Sosa et al., 2012; Morrison et al., 2013).

Bei den inneren Patienten ist häufig die körperliche Leistungsfähigkeit reduziert (Graf, & Halle, 2007, S. 324) und Personen mit orthopädischen Erkrankungen sind häufig durch Beschwerden am Bewegungsapparat eingeschränkt. Die Gruppe mit neurologischen Erkrankungen leidet häufig an ausgeprägten motorischen Störungen wie z. B. Gleichgewichtsproblemen, Einseitigkeit, Wahrnehmungsstörungen.

Erwartungsgemäß zeigte sich, dass Personen mit körperlichen Einschränkungen (Neurologie, Innere, Orthopädie) langsamer waren als wenig motorisch Eingeschränkte (HNO, Gesunde, Sportstudierende) (s. Abb. 28, S. 98). Für weiterführende Berechnungen wurden dementsprechend die Gesunden und HNO-Patienten gemeinsam als „Körperlich Gesunde“ und die Neurologie-, Innere- und Orthopädiepatienten als „Körperlich Kranke“ betrachtet.

Obwohl beide Geschlechter bei der Lösungshäufigkeit annähernd gleich abschnitten (Tab. 26, S. 95), benötigten mit Ausnahme der Sportstudierenden die Männer weniger Zeit als die Frauen, wenn alle Items gelöst wurden (s. Tab. 28, S. 97). Beim klassischen TUG konnten solche geschlechtsabhängige Unterschiede von Ibrahim, Singh und Shahar (2017) nicht bestätigt werden. Dieser Unterschied ist möglicherweise auf die höheren Kraftbeanspruchungen beim TUG<sub>plus</sub> (z. B. Kasten umstoßen, Kasten tragen) zurückzuführen.

Die Ergebnisse zeigten weiterhin eine altersabhängige Zunahme der Testzeit bei der Gesamtstichprobe (Anstieg um 2 bis 3 s; s. Tab. 29, S. 99; Abb. 29, S. 100). Dies entspricht, wenn auch bei anderen Altersspannen, den Studienergebnissen zum TUG von Steffen, Hacker und Mollinger (2002) und der Metaanalyse von Bohannon (2006; s. Tab. 8, S. 48). Dass dementsprechend Einschränkungen bei Alltagstätigkeiten auftreten, konnte auch von Shumway-Cook et al. (2005) und später im Deutschen Alterssurvey (2008) gezeigt werden (s. Tab. 1, S. 9). Als Ursachen werden neben krankheitsbedingten Limitationen häufig ein sich verschlechterndes Verhältnis von Muskel- und Fettmasse (Hunter et al., 2001; Janssen, & Ross, 2005; Kehayias et al., 1997; Visser et al.,

2005), eine damit verbundene geringere Kraft (Frontera et al., 1991; Sallinen et al., 2010) und vermehrte Koordinationsprobleme (Cumming et al., 2000) genannt.

Hinsichtlich der Personen- bzw. Patientengruppen ist eine altersabhängige Zunahme der Testzeit bei neurologischen und HNO-Patienten sowie bei gesunden Personen zu erkennen (s. Abb. 30, 31, 34, S. 101 ff.). Bei orthopädischen Patienten ist erst bei Personen im Alter ab 66 Jahren ein Anstieg zu verzeichnen (s. Abb. 33, S. 104).

Vergleicht man die Testzeit (s) mit der Ausführungsökonomie (Punktwert), besteht ein negativer Zusammenhang (s. Abb. 35, S. 106). Je höher die erreichte Punktzahl (hohe Ausführungsökonomie), desto schneller wurde der Parcours durchschnittlich durchlaufen.

Ausnahmen bestätigen allerdings hier die Regel. Die Einzelfallbetrachtungen zeigen, dass diese Annahme nicht immer zutraf. Eine Probandin erreichte trotz mäßiger Bewegungsökonomie (6 Punkte) eine etwas oberhalb dem Mittel (22,1 s) liegende Testzeit von 23,8 Sekunden, während eine zweite Probandin trotz optimaler Bewegungsausführung (11 Punkte) mit rund 56 Sekunden recht lange für den Parcours benötigte (s. Abb. 35, S. 106). Für die hohe Testzeit kann die Krebserkrankung aber auch das im Vergleich zur Gesamtstichprobe relativ hohe Alter der Teilnehmerin (75 Jahre) verantwortlich sein, da bei der Therapie von Krebserkrankungen gerade bei Frauen als Nebenwirkungen Osteoporose auftreten kann und in dem Fall Angst vor Stürzen und Bewegungsunsicherheit eine nicht unerhebliche Rolle spielen.

Die bei Probanden mit beispielsweise Herzkreislaufkrankungen notwendige Vorselektion durch eine ärztliche Untersuchung verringerte das Risiko, dass beim Test kardiale Zwischen- bzw. Notfälle auftraten (ACSM, Thompson et al., 2013). Dennoch wurden auch wenig leistungsfähige Patienten in den Test integriert, um auch diesen Grenzbereich zu überprüfen.

Dazu wurden Personen mit höhergradigen körperlichen Einschränkungen analysiert, die eine Testzeit über 48 Sekunden erreichten (s. Tab. 30, S. 107). Ähnlich wie bei der Ausführungsökonomie (s. Tab. 24, S. 93), hat die Krankheitsschwere auch einen starken negativen Einfluss auf die Testzeit.

Die aus den insgesamt aufgeführten Erkrankungen zusätzlich resultierenden psychischen Auswirkungen, wie Vermeidung schmerzauslösender Bewegungen und/oder Bewegungsängste, können die Lösungshäu-

figkeit der Items und die Testzeit ebenfalls beeinflusst haben (Pfingsten, 2004).

Haben Patienten hieraus Kompensationsstrategien entwickelt, sind die eigentlichen motorischen Defizite unter Umständen nicht mehr so offensichtlich und die Probanden können den Parcours trotz ihrer Einschränkungen vollständig bewältigen (Smith et al., 1996).

Die Gruppen, die körperlich/konditionell in der Lage sind lediglich submaximale Leistungen bzw. Geschwindigkeiten beim TUG<sub>plus</sub> zu erbringen (z. B. wenig belastbare Herz-, COPD-Patienten), können natürlich trotzdem im Hinblick auf ihre motorische Leistungsfähigkeit qualitativ bewertet werden – ohne Relevanz der Testzeit (Lösungshäufigkeit der Items bzw. Ausführungsökonomie, s. Tab. 14, S. 66).

Um die Ökonomie des Verfahrens, insbesondere für die Testleiter, zu gewährleisten, wurde bei dieser Untersuchung aufgrund der Korrelation zwischen Testzeit und Ausführungsökonomie (s. Abb. 35, S. 106) und der einfacheren Bewertung für weitere Berechnungen die Testzeit als alleiniger Messparameter verwendet.

Schlussfolgernd wird folgendes festgehalten:

- Im Gegensatz zu den Befunden des TUG absolvieren Männer den TUG<sub>plus</sub> schneller als Frauen, was möglicherweise auf die Kraftanteile (z. B. Kasten umstoßen) zurückzuführen ist, die für Männer und Frauen nicht differenziert wurden. Dies gilt nicht für die Gruppe der Sportstudierenden. **Hypothese 3** (S. 91) ist daher bedingt anzunehmen.
- Personen mit körperlichen Einschränkungen (Neurologie, Innere, Orthopädie) benötigen mehr Zeit als körperlich gesunde Personen und die Vergleichsgruppe der Sportstudierenden. Die Zeiten körperlich Gesunder und HNO-Patienten unterscheiden sich nicht. **Hypothese 4** (S. 91) ist nur bedingt anzunehmen.
- Mit zunehmendem Alter nimmt auch die Testzeit durchschnittlich zu. Dies trifft nicht auf jede Patientengruppe zu. So sind bei Patienten mit inneren Erkrankungen keine altersabhängigen Testzeiten festzustellen. Desweiteren werden die Unterschiede häufig erst im höheren Alter sichtbar. **Hypothese 5** (S. 91) kann also nur bedingt angenommen werden.
- Eine ökonomische Ausführung der Items steht in Zusammenhang mit einer geringen Testzeit. **Hypothese 6** (S. 91) kann somit an-

genommen werden. Bei den folgenden Berechnungen wurde dementsprechend nur die Testzeit berücksichtigt. In der Praxis sind die aufgeführten Aspekte zur Beurteilung der Bewegungshandlungen (s. Tab. 14, S. 66) für eine differenzierte Beurteilung dennoch sinnvoll.

### 6.1.3 Reliabilität

Hinsichtlich der Reliabilität wurden interne Konsistenz und Paralleltest-Reliabilität überprüft.

Die interne Konsistenz wurde geschlechtsunabhängig und -abhängig geprüft (s. Kap. 5.4.1, S. 108). Hinsichtlich der Itemzeiten zeigten die Items bei Männern und Frauen mittlere bis hohe Korrelationen zueinander (s. Tab. 34 - 36, S. 111 f.). Mit einem  $\alpha = 0,85$  (Männer 0,89, Frauen 0,79) ist die von Schmitt (1996, S. 351) definierte Mindestanforderung an die interne Konsistenz gegeben ( $\alpha = 0,70$ ). Sie kann als gut beurteilt werden und ist vergleichbar mit der Berg-Balance-Scale ( $\alpha = 0,83$ ; Berg et al., 1995) und dem Six-Minute Walk Test ( $\alpha = 0,87$ ; Collins et al., 2008).

Die Trennschärfekoeffizienten der Items sind ebenfalls zufriedenstellend (s. Tab. 37 - 39, S. 113 f.). Die Aufgaben können als homogen angesehen werden (Lienert, & Raatz, 1998, S. 202). Lediglich das Weglassen des Items „Aufstehen (Bank), Gehen“ würde bei Frauen eine minimale Erhöhung der internen Konsistenz bewirken (s. Tab. 39, S. 114). Für dieses Item sind also Alternativen denkbar, die die Ökonomie des Tests ohne Qualitätsverlust weiter erhöhen. So könnte der Start anstatt auf einer Turnbank beispielsweise aus dem Hochstart erfolgen oder erst beim Item „Hindernis überwinden“ beginnen.

Die benötigten Zeiten für jedes einzelne Item waren Datenbasis für die Berechnung der internen Konsistenz. Im Vorhinein wurden hierzu Start- und Endpunkte für die jeweiligen Zeitmessungen definiert (s. Tab. 32, S. 110). Bei Wiederholungsmessungen müssen demnach diese Start- und Endpunkte genau eingehalten werden, um die Vergleichbarkeit der Messungen zu gewährleisten.

In einem weiteren Schritt wurde die Paralleltest-Reliabilität überprüft. Da sich diese Methode vor allem bei Tests eignet, bei denen eine Schnelligkeitsleistung gemessen wird (Lienert, & Raatz, 1998, S. 213) und der Paralleltest (umgekehrter Parcoursdurchlauf) einfach zu konzipieren war, eignete sich diese Methode als weiterer Reliabilitätsmarker.

Die Zeiten der klassischen und umgekehrten Parcoursreihenfolge korrelierten hoch. Ebenfalls hohe Zusammenhänge zeigten sich hinsichtlich des subjektiven Anstrengungsempfindens (s. Tab. 41, S. 117). Die benötigten Zeiten sind zudem unabhängig von Geschlecht, Alter, BMI und Erkrankung und somit unabhängig von der Anordnung der Items (s. Kap. 5.4.2.5, S. 117).

Die Parcoursdurchläufe in klassischer und umgekehrter Reihenfolge wurden unmittelbar aufeinander folgend durchgeführt. Störeffekte durch Zeiteinflüsse, die die interne Validität einschränken, konnten somit ausgeschaltet werden.

Die Probanden wurden randomisiert den vier möglichen Testreihenfolgen zugeteilt. Der Einfluss, ob der Parcours zuerst in vorgesehener oder umgekehrter Itemreihenfolge durchgeführt wurde, konnte so kontrolliert werden.

Schlussfolgernd ist festzuhalten:

- Die Items des TUG<sub>plus</sub> zeigen hohe Zusammenhänge zueinander (gemessen über Einzelzeiten) und somit eine gute interne Konsistenz. Die **Hypothese 7** (S. 110) kann somit angenommen werden.
- Die Items zeigen Trennschärfekoeffizienten  $r_{it} > 0,6$ . **Hypothese 8** (S. 110) kann somit angenommen werden.
- Die Testzeiten bei vorgesehener und umgekehrter Itemanordnung korrelieren hoch. Diese unterscheiden sich nicht hinsichtlich Alter, Geschlecht und BMI und sind damit unabhängig von der Parcoursreihenfolge. Die unterschiedliche Anordnung wird auch nicht subjektiv anstrengender oder weniger anstrengend empfunden. Somit können die **Hypothesen 9** und **10** (S. 116) angenommen werden.

#### 6.1.4 Responsivität

Bei neurologischen und HNO-Patienten zeigten sich bei direkt aufeinander folgenden Durchgängen (Vortest 1, 2) hohe Zusammenhänge hinsichtlich der Testzeit (s. Tab. 43, S. 121). Testeffekte, die gerade bei wiederholten Testungen motorischer Aufgaben vorkommen (Boutin, Panzer, & Blandin, 2013), können somit weitestgehend ausgeschlossen werden (Kaczmarek, Schwarz, Backes, & Wydra, 2016b, S. 143; Kaczmarek, Schwarz, & Wydra, 2017a, S. 149).

Bei der Überprüfung der Responsivität ergaben sich nach einem zweiwöchigen Treatment Leistungsverbesserungen (Vortest 2 zu Nachtest) bei beiden Patientengruppen (s. Abb. 39, S. 122). Hierbei muss ange-



merkt werden, dass die interventionsbedingten Zeiteinflüsse wiederum die Prüfung der Retest-Reliabilität stören. Zu folgern, dass die Leistungsveränderungen primär auf das Treatment zurückzuführen sind, ist keineswegs zwingend. So zeigen sich bereits anhand der Prüfung der Testeffekte zwar geringe, aber dennoch vorhandene Wiederholungseffekte. Der erste Testdurchlauf kann somit neben der Intervention das Ergebnis der Abschlussmessung beeinflusst haben. Welche Therapieformen für die Leistungssteigerungen verantwortlich sind, konnte in dieser Untersuchung aufgrund nicht standardisierter Therapieprogramme ebenfalls nicht beantwortet werden (Kaczmarek et al., 2016b, S. 143; Kaczmarek et al., 2017a).

Schlussfolgernd wird festgehalten:

- Das Verfahren führt bei Patienten der Neurologie und HNO zu vernachlässigenden Testeffekten. Nach zweiwöchiger Intervention zeigen sich Leistungsveränderungen in Form geringerer Testzeiten. Die **Hypothesen 11** und **12** (S. 121) sind demnach bestätigt.

#### 6.1.5 Validität

Nach Testung der Reliabilität sollte die Kriteriumsvalidität überprüft werden (s. Kap. 5.6.1, S. 123). Hierzu dienten eine fahrradergometrische Untersuchung (W/kg KG), das subjektive Anstrengungsempfinden (Borg-Skala), der TUG (Testzeit), die MBD (Summenscore), der MFI (Summenscore) und die Handkraft (kg) als Kriterien (s. Kap. 5.1.3, S. 72; Kaczmarek et al., 2017a, S. 143).

Bis auf TUG und Handkraft (links), MBD und Handkraft (links, rechts) korrelierten die Tests untereinander, was die Verwendung als Validitätskriterien ermöglichte (s. Tab. 47, S. 126).

Die auf dem Fahrradergometer erreichte Leistung korrelierte am stärksten mit der Testzeit des TUG<sub>plus</sub> (s. Tab. 48, S. 126). Die Teilnehmer erreichten bei Ausbelastung eine durchschnittlich maximale Leistung von 2,4 W/kg KG. Bei abnehmender Leistungsfähigkeit erhöhte sich auch die Testzeit der Probanden. Dass Mobilitätseinschränkungen, insbesondere Gehprobleme, mit Ausdauerleistungen korrelieren, wiesen bereits Simonsick et al. (2008) und Dallmeijer und Van der Woude (2001) nach.

Der klassische TUG wurde mittels Barthel-Index ( $r = -0,51$ ) und Berg-Balance-Scale validiert ( $r = -0,81$ ) validiert. Dementsprechend konnte der TUG<sub>plus</sub> in Bezug zum TUG gesetzt werden ( $r = 0,64$ ; Tab. 48, S. 126).

Die bei der MBD getesteten motorischen Fähigkeiten zeigten in der Summe mittlere Korrelationen zum TUG<sub>plus</sub>.

Weiterhin zeigt der TUG<sub>plus</sub> eine mittlere Korrelation ( $r = -0,38$ ) zum MFI. Je schneller die Personen den Parcours durchliefen, desto weniger Alltagstätigkeiten waren anhand des MFI eingeschränkt.

Die Ergebnisse zeigen auch, dass die Handkraft mit der benötigten Zeit beim TUG<sub>plus</sub> korreliert. Dies bestätigt die Ergebnisse von Rantanen et al. (1999). Sie wiesen nach, dass eine schwächere Handkraft mit Problemen bei der Hausarbeit, beim Treppensteigen, beim Aufstehen von einem Stuhl, beim Tragen schwerer Gegenstände und mit geringeren Gehgeschwindigkeiten korreliert.

Da in der durchgeführten Untersuchung keine Einzelzeiten der Items erfasst wurden, können keine Rückschlüsse auf Zusammenhänge zu bestimmten, beim TUG<sub>plus</sub> integrierten Bewegungsformen, gezogen werden. Im Gegensatz dazu stellten verschiedene Studien direkte Zusammenhänge zwischen bestimmten motorischen Bewegungen und Activities of Daily Living her (Al Snih, Markides, Ottenbacher, & Raji, 2004; Rantanen, Era, & Heikkinen, 1994; Rantanen et al., 1999). Al Snih et al. (2004) verwendeten hierbei Beurteilungssysteme, die eine qualitative Erfassung (über Skalen) von Bewegungen bzw. Funktionalitäten beinhalten.

Die mittleren Zusammenhänge zu den externen Kriterien (TUG, MBD, MFI, Handkraft, Fahrradergometrie) bestätigen die Validität des Screenings.

Die Hypothese, dass die Items des TUG<sub>plus</sub> tatsächlich eine einfaktorielle Struktur besitzen, wurde mittels Faktorenanalyse überprüft (s. Kap. 5.6.2, S. 127). Die explorativen und konfirmatorischen Faktorenanalysen führten im Gesamtmodell sowie in den Modellen der „körperlich Gesunden“ und „körperlich Kranken“ zu einer jeweils einfaktoriellen Struktur mit guter Anpassung (Kaczmarek et al., 2017b). Der gefundene Faktor wurde ICF-orientiert als „Mobilität“ bezeichnet (s. Abb. 41, S. 134 f.). Somit kann die Konstruktvalidität des TUG<sub>plus</sub> bestätigt werden. Die einfaktorielle Struktur bestätigt den Charakter eines Screenings und erlaubt im Nachgang die Berechnung von Trennpunkten (Cut-Off-Werte).

Zur Überprüfung der Konstruktvalidität wurde eine Dichotomisierung der Probanden in „Körperlich Krank“ bzw. „Körperlich Gesund“ gewählt – trotz der Einteilungsproblematik in zwei Extremgruppen – da diese Be-

dingung die Voraussetzung für die weitere Bestimmung von Cut-Off-Werten ist.

Folgende Schlussfolgerungen werden formuliert:

- Die Außenkriterien (Fahrradergometrie, TUG, MBD, MFI, Handkraft) eignen sich aufgrund mittlerer Korrelationen zueinander zur Bestimmung der Kriteriumsvalidität. **Hypothese 13** (S. 125) kann bestätigt werden.
- Der TUG<sub>plus</sub> zeigt mittlere Korrelationen zur fahrradergometrischen Leistung (W/kg KG), MBD (Summenwert), Handkraft (kg), zum TUG (s) und MFI (Summenwert). Die Validität ist somit bestätigt. **Hypothese 14** (S. 125) kann ebenfalls angenommen werden.
- Die Items zeigen im Gesamtmodell sowie bei „körperlich Gesunden“ und „körperlich Kranken“ jeweils eine einfaktorielle Faktorenstruktur mit einer guten Modellanpassung. Der gefundene Faktor kann ICF-orientiert als „Mobilität“ bezeichnet werden. **Hypothese 15** (S. 128) wird angenommen.

## 6.1.6 Grenzwertbestimmung

### 6.1.6.1 Ergebnisverteilung

Zur Bestimmung eines Normbereiches beim TUG<sub>plus</sub> wurde die Verteilung der Testzeiten der Gesamtstichprobe beurteilt (s. Kap. 5.7.1, S. 136). Hierzu dienten rein deskriptive Maße (Quartile, Median).

Wird das obere Quartil herangezogen, liegt ein möglicher Trennpunkt bei rund 27 Sekunden (s. Abb. 42, S. 137). Personen, die mehr als 27 Sekunden für den Parcours benötigen, könnten als motorisch auffällig charakterisiert werden. Dieser Trennpunkt ist zunächst unabhängig von einem sogenannten Gold- bzw. Referenzstandard (vgl. Kap. 3.3.3, S. 37). So werden bei der Grenzwertfestlegung keine Personenmerkmale wie z. B. Gesundheitszustand erfasst. Auch können dadurch keine speziellen Fragestellungen beantwortet werden, wie dies beispielsweise anhand von diagnostischen Studien zum TUG aufgezeigt wurde (vgl. Tab. 11, S. 55).

Um einen Trennpunkt anhand eines Referenzstandards zu bestimmen und spezielle Fragestellungen beantworten zu können, wurde im Weiteren die diagnostische Güte überprüft.

#### 6.1.6.2 Diagnostische Güte

Anhand eines Screenings soll eine Grobbeurteilung eines Merkmals vorgenommen werden. Dies erfordert die Festlegung eines Grenzwertes (Cut-Off-Wert; s. Kap. 3.3.3, S. 37). Zur Bestimmung von Cut-Off-Werten wurden zwei Referenzstandards verwendet, die jeweils eine Merkmal-Dichotomisierung in zwei Extremwerte zuließen (s. Kap. 3.3, S. 32).

Durch Zusammenfügung beider Standards konnte die Trennschärfe bei kombinierten Einschränkungen überprüft werden (s. Kap. 5.7.2, S. 138).

##### *Gesundheitsstatus*

Häufig wird die diagnostische Güte eines Testverfahrens an der Vorhersagekraft eines körperlichen Zustandes von Personen gemessen (vgl. Tab. 11, S. 55). So wurde in einer Studie zum TUG der Gesundheitsstatus (gesund/krank) als Referenzstandard benutzt (Bischoff et al., 2003).

Vergleichbar sollte hierzu überprüft werden, ob anhand des Screenings Personen mit körperlichen Einschränkungen identifiziert werden können, die tatsächlich Probleme hinsichtlich Motorik bzw. Mobilität haben. Es zeigte sich, dass sich die dichotomisierten Personengruppen hinsichtlich der Testzeit deutlich unterscheiden. Personen ohne körperliche Einschränkungen durchliefen den Parcours rund 7 Sekunden schneller als Personen mit Beeinträchtigungen (s. Abb. 44, S. 144).

Der Vergleich der Perzentile und die weite Spannweite des 95%-Konfidenzintervalls (17 s; 26 s) machte eine genauere Analyse hinsichtlich diagnostischer Qualitätskriterien erforderlich.

Anhand des Youden-Index von 0,41 ergibt sich ein Cut-Off-Wert bei einer Testzeit von 21 Sekunden, bei der 69 % (Sensitivität) der Personen mit körperlichen Beeinträchtigungen gut identifiziert werden können (s. Tab. 58, S. 145; Fluss et al., 2005). Zudem ist eine gute Balance der Parameter Sensitivität (0,69) und Spezifität (0,72) gegeben, was als akzeptable diagnostische Güte gewertet werden kann (Perkins, & Schisterman, 2006).

Werden Trennpunkte kleiner 21 Sekunden gewählt (höhere Sensitivität, geringere Spezifität), kann entsprechend ein höherer Prozentsatz an Personen mit körperlichen Einschränkungen erkannt werden, als bei Trennpunkten größer 21 Sekunden (geringere Sensitivität, höhere Spezifität). Entfernt man sich also von diesem gut ausbalancierten Cut-Off-Wert von 21 Sekunden, werden mehr Personen falsch klassifiziert.

Im Vergleich hierzu, zeigt der TUG teils geringere (Haines et al., 2008; Kojima et al., 2015; Kristensen et al., 2007; Nordin et al., 2008; Savva et al., 2012; Shimada et al., 2009; Whitney et al., 2004) und teils höhere Indizes (Sensitivität, Spezifität, Youden-Index; Shumway-Cook et al., 2000; Rose et al., 2002; Chiu et al., 2003) hinsichtlich der Fähigkeit, sturzgefährdete, d. h. motorisch auffällige Personen zu erkennen (vgl. Tab. 11, S. 55).

Zusätzlich diente der Ratz-Index zur Überprüfung der Trefferquote des Screenings. Bei dem veranschlagten Cut-Off-Wert ergibt sich eine moderate Verbesserung im Vergleich zum Zufall (Ratz = 0,47; s. Tab. 58, S. 145). Bei einem Trennpunkt von 16 Sekunden wäre mit einem Ratz-Index von 0,66 eine hohe Verbesserung möglich (s. Tab. 58, S. 145; Loeber, & Dishion, 1983; Marx, & Lenhard, 2010). Allerdings liegt dann die Sensitivität bei 94 % und die Spezifität bei lediglich 38 %, was bedeutet, dass zwar ein hoher Prozentsatz an Personen mit körperlichen Beeinträchtigungen richtig identifiziert werden kann, aber nur eine relativ niedrige Anzahl an Personen ohne Beeinträchtigungen richtig klassifiziert wird.

Als weitere Qualitätskriterien liegt die ROC-Kurve weit entfernt von der Winkelhalbierenden (s. Abb. 47, S. 150) und der AUC-Wert (0,67) in einem wünschenswerten Bereich zwischen 0,5 und 1 (s. Tab. 58, S. 145; Hanley, & McNeil, 1982). Dies bedeutet, dass anhand unterschiedlicher Cut-Off-Werte eine deutlich höhere Trefferquote wie 50 % erzielt werden können.

Die Wahrscheinlichkeitsverhältnisse führten zu einer moderaten diagnostischen Evidenz (s. Kap. 5.7.2.5.1, S. 146; Jaeschke, Guyatt, & Lijmer, 2002). Für Personen mit körperlichen Einschränkungen ist die Wahrscheinlichkeit ein positives TUG<sub>plus</sub>-Ergebnis zu erzielen um das Zweieinhalbfache höher als für Personen ohne körperliche Beeinträchtigungen (LR+ = 2,5). Umgekehrt ist für Personen mit körperlichen Beeinträchtigungen die Wahrscheinlichkeit ein negatives Ergebnis zu erzielen 60 % geringer als für Personen ohne solche Einschränkungen (LR- = 0,4). Diese Verhältnisse zeigen eine bessere LR+ aber eine schlechtere LR- als Kristensen et al. (2007) für den TUG gefunden haben, um eine erhöhte Sturzgefahr innerhalb von sechs Monaten nach Hüftfraktur zu prognostizieren (LR+ = 1,5, LR- = 0,1).

Weiterhin zeigten Personen, die den TUG<sub>plus</sub> in mehr als 21 Sekunden absolvierten, Probleme hinsichtlich Mobilität (TUG) und motorischer Fähigkeiten (MBD; s. Abb. 45, S. 146; Tab. 59, S. 147). Dies ist ebenfalls mit Ergebnissen aus TUG-Validitätsstudien vergleichbar. Dort zeig-

te sich, dass je niedriger die Geschwindigkeit, desto mehr Mobilitäts- und Gleichgewichtsprobleme können mittels Tinetti-Test und Berg-Balance-Scale erkannt werden (Podsiadlo, & Richardson, 1991; Bennie et al., 2003).

Die Berechnungen der diagnostischen Güte sind jedoch mit Limitationen verbunden. So ist auch der berechnete Cut-Off-Wert nicht allgemeingültig, sondern zunächst nur auf die getesteten Personengruppen anwendbar. Bei anderen Personen- bzw. Patientengruppen mit anderen klinischen Krankheitsschweren, muss der Cut-Off-Wert angepasst bzw. neu berechnet werden (Spektrum-Bias; Knottnerus et al., 2009). Diesbezüglich konnte in einem Review gezeigt werden, dass je schwerer die Erkrankung ist, desto höher fällt die Sensitivität eines diagnostischen Instruments aus (Whiting, Rutjes, Reitsma, Glas, Bossuyt, & Kleijnen, 2004).

Des Weiteren limitieren der gewählte Referenzstandard und der spätere praktische Nutzen das Setzen eines Cut-Off-Wertes. In einem Review haben Reitsma et al. (2009) Methoden zum Umgang mit solchen „Referenzstandards“ zusammengetragen. Es zeigte sich, dass in manchen Studien die Genauigkeit des Screenings anhand eines Vergleiches mit Ergebnissen des Referenzstandards oder anderer Tests beurteilt wird. Um die Exaktheit des Standards zu erhöhen, werden zudem häufig zwei oder mehrere Referenzstandards verwendet. Hierbei werden die Testergebnisse kombiniert und zu einem Gesamt-Referenz-Ergebnis zusammengefasst. So wurde in einer weiteren Berechnung der zusätzliche Referenzstandard „Motorische Leistungsfähigkeit“ verwendet, der zunächst separat und anschließend mit dem „Gesundheitsstatus“ zu einem Standard zusammengeführt wurde.

### *Motorische Leistungsfähigkeit*

Weiterhin sollte die diagnostische Güte anhand der Stufengeschwindigkeit 7 km/h bestimmt werden. Ziel war es Personen, die motorisch laufen können von Personen, die motorisch nicht laufen können, zu unterscheiden. Diese Geschwindigkeitsstufe wurde gewählt, um möglichst konditionelle Aspekte als Abbruchgrund auszuschließen.

Personen, die diese Geschwindigkeit schafften, benötigten durchschnittlich weniger Zeit beim TUG<sub>plus</sub> als diejenigen, die dazu nicht fähig waren (s. Abb. 46, S. 148). Mittels Youden-Index liegt der optimale Trennpunkt bei 25 Sekunden (s. Tab. 60, S. 148; Kaczmarek, Schwarz, & Wydra, 2017c). Aufgrund höherer Spezifitätswerte können hier mehr Personen, die das Kriterium erfüllen als solche erkannt werden. Bei geringeren Testzeiten

(< 25 s) kann ein höherer Prozentsatz von Personen mit Mobilitätsstörungen erkannt werden als bei höheren Trennpunkten. Ob dieses Kriterium und somit die Testzeit (TUG<sub>plus</sub>) von 25 Sekunden tatsächlich als Prädiktor von Mobilitätsproblemen angesehen werden kann, bleibt in dieser Untersuchung offen. Häufig wurde in Studien hierzu die Fähigkeit „Gehen können“ als „Minimalnorm“ (Israel, 1983) verwendet. So führten Guralnik et al. (2000) eine diagnostische Untersuchung durch, in der die Ganggeschwindigkeit als Prädiktor für Mobilitätsprobleme überprüft wurde. Anhand von ROC-Analysen konnten sie nachweisen, dass Personen, die weniger als 1 m/s (3,6 km/h) gehen können vermehrt Mobilitätsprobleme aufweisen. Cesari et al. (2005) bestätigten später diesen Cut-Off-Wert.

### *Motorische Leistungsfähigkeit und Gesundheitsstatus*

Durch die Zusammenführung beider Standards (Gesundheitsstatus: n = 584; Motorische Leistungsfähigkeit: n = 119) konnte der Referenzstandard verschärft werden, was allerdings zwangsläufig zu einer kleineren aber für die weiteren Berechnungen ausreichenden Schnittmenge führte (n = 65).

Anhand des Youden-Index fand sich der gleiche Cut-Off-Wert (21 s) wie beim Referenzstandard „Gesundheitsstatus“, allerdings mit einer besseren diagnostischen Güte (s. Kap. 5.7.2.5.3, S. 149). Im Vergleich zu den vorherigen Berechnungen ist die Trefferquote hier deutlich größer (Ratz-Index = 0,82). Zudem besteht eine 88-prozentige Wahrscheinlichkeit, dass anhand positiver Testresultate tatsächlich Personen mit körperlichen Einschränkungen identifiziert werden können (Sensitivität; Zhou et al., 2002).

Auch die Wahrscheinlichkeitsverhältnisse (Likelihood Ratio) zeigen höhere Werte als die separaten Referenzstandards (LR+ = 7,3, LR- = 0,1; Jaeschke et al., 2002). Die Wahrscheinlichkeit für körperlich beeinträchtigte Personen ein positives Ergebnis zu erzielen ist 7,3 mal höher als für körperlich Unbeeinträchtigte (LR+ = 7,3). Im Gegensatz dazu ist die Wahrscheinlichkeit für körperlich Beeinträchtigte im TUG<sub>plus</sub> ein negatives Ergebnis zu erzielen 90 % niedriger als für körperlich Unbeeinträchtigte (LR- = 0,1). In einer Metaanalyse wurden zum klassischen TUG die Wahrscheinlichkeitsverhältnisse aus zehn Studien analysiert (Barry, Galvin, Keogh, Horgan, & Fahey, 2014). Alle Studien zusammengefasst zeigen eine deutlich geringere Testeffizienz (LR+ 1,2; LR- 0,2) als die vorliegenden Ergebnisse.

Der AUC-Wert liegt bei einer Zeit von 21 Sekunden bei 0,85, was ein sehr gutes diagnostisches Gütemaß darstellt (Hanley, & McNeil, 1982; Abb. 47, S. 150).

Die Bestimmungen der diagnostischen Güte verdeutlichen die in Kapitel 3.3.4 (S. 41) erläuterte Grenzwertproblematik. Bei welcher Testzeit das Ergebnis als „normal“ charakterisiert werden kann (Israel, 1983, S. 360), ist hier vom gewählten Referenzstandard abhängig (Reitsma et al., 2009, S. 799). Zudem sollte bei der Wahl des Cut-Off-Wertes die jeweilige Zielstellung berücksichtigt werden (Fokus auf Sensitivität oder Spezifität; Lalkhen, & McCluskey, 2008).

Die gesetzten Trennpunkte können als „Minimalnorm“ (Israel, 1983; s. Kap. 3.3.4, S. 41) verstanden werden. Werden diese überschritten, ist dies eine Indikation für eine motorische Intervention.

Die wesentlichen Schlussfolgerungen aus diesem Kapitel lauten wie folgt:

- Ähnlich wie beim TUG konnten auch beim TUG<sub>plus</sub> Cut-Off-Werte identifiziert werden. Die Berechnungen verdeutlichen das Problem der Dichotomisierung der TUG<sub>plus</sub>-Ergebnisse.
- Die Untersuchungen zur diagnostischen Güte verdeutlichen, dass die entsprechenden Cut-Off-Werte in hohem Maße vom gewählten Referenzstandard abhängen.
- Referenzstandard „Gesundheitsstatus“: Bei einem Cut-Off-Wert von 21 Sekunden können körperlich beeinträchtigte und unbeeinträchtigte Personen bei einer akzeptablen diagnostischen Güte (69 % Sensitivität, 72 % Spezifität), unterschieden werden. **Hypothese 16** (S. 143) kann ebenfalls angenommen werden.  
Für Personen, die mehr als 21 Sekunden für den Parcours benötigen, sind körperliche Einschränkungen und Probleme hinsichtlich Mobilität und motorische Fähigkeiten zu erwarten. Somit ist **Hypothese 17** (S. 143) anzunehmen.
- Referenzstandard „Motorische Leistungsfähigkeit“: Bei einem Trennpunkt von 25 Sekunden weist das Screening anhand dieses Kriteriums eine annehmbare diagnostische Güte auf. **Hypothese 18** (S. 143) ist somit ebenfalls anzunehmen.
- Die Zusammenfügung beider Kriterien zu einem Referenzstandard „Motorische Leistungsfähigkeit“ und „Gesundheitsstatus“ führt zu einem Cut-Off-Wert von 21 Sekunden mit der besten diagnostischen Güte. **Hypothese 19** (S. 143) ist bestätigt.



Um die Aussagesicherheit von Normwerten deutlich zu erhöhen, wird daher empfohlen als Bezugsgröße zwei zueinander passende Referenzstandards zu wählen.

## 6.2 Methodenkritik

Während der Untersuchungsphase sind methodische Besonderheiten aufgetreten, die im Vorhinein nicht erwartet werden konnten, deren Berücksichtigung aber für zukünftige Studien zu der Thematik hilfreich sein können.

Der TUG<sub>plus</sub> hat sich im Vergleich zum klassischen TUG als motorisch differenzierter und aussagekräftiger erwiesen, ist aber mit einem zeitlich und materiell etwas höheren Aufwand verbunden. Für sehr wenig Belastbare oder Hochaltrige (Testzeiten  $\geq 48$  s) erscheint der ursprüngliche TUG mit den aktualisierten Normbereichen – v. a. auch aus ökonomischer Sicht – ausreichend (s. Tab. 30, S. 107).

Auch bei jüngeren bzw. leistungsfähigeren Personen ist die reine Testzeit des Screenings mit rund 22 s dreimal so lange wie beim TUG (ca. 7 s). Die aufwändigere Testinstruktion verlängert diese Zeit zusätzlich.

Aufgrund der fünf komplexeren Items ist im Gegensatz zum TUG zudem ein Vorüben notwendig, um zum einen die verschiedenen Items kennenzulernen und zum anderen passende Lösungsstrategien zu erproben. Zudem begünstigt jedes zusätzlich aufgenommene Item den Einfluss weiterer Störfaktoren auf das Testergebnis.

Die mit dem Messparameter Zeit verbundenen Einflussfaktoren Testinstruktion, Startkommando, Animation während des Tests und Motivation sind bedeutsame Störquellen, die im Vergleich zum TUG auch durch die Hinzunahme weiterer Items erhöht werden. Hierbei kann die Testinstruktion als Hauptstörgröße genannt werden. So zeigte sich bereits in Studien zum TUG, dass unterschiedliche Instruktionen („schnellstmöglich“, „schnellstmöglich und sicher“, „gewohntes Gehtempo“, „selbstgewähltes Gehtempo“, keine Anweisung) die Testzeit maßgeblich beeinflussen (s. Tab. 9, S. 51; Maanum et al., 2010; Pertille et al., 2012; Granacher et al., 2013; Klima et al., 2013; Lee et al., 2013; Sakurai et al., 2013; Regterschot et al., 2014; Kumar et al., 2015; Mulasso et al., 2015). Um bei inter- und intraindividuellen Vergleichen aussagekräftige Ergebnisse zu erhalten und das maximal mögliche Tempo der Probanden zu erfassen, soll der TUG<sub>plus</sub> in Anlehnung an TUG-Studien von Pertille et al. (2012) und Sakurai et al. (2013) „schnellstmöglich“ durch-

laufen werden, was eine hohe Motivation impliziert (Whitney, Lord, & Close, 2005). Da sich auch bereits in Vorversuchen zum klassischen TUG je nach Instruktion („schnellstmöglich“ vs. „schnellstmöglich und sicher“) erhebliche Unterschiede hinsichtlich der Testzeit ergaben, wurde diese Instruktion gewählt. Dadurch wurde bei den im Rahmen dieser Arbeit durchgeführten Untersuchungen keine erhöhte Sturzhäufigkeit protokolliert ( $< 1\%$ ).

Die Testökonomie ist mit nur einem Parameter (Zeit) indiskutabel hoch. Ob dieser einzelne Befund einem Therapeuten allerdings ausreichend erscheint, um Aussagen über die Mobilität einer Person zu treffen, ist fraglich. Bei insgesamt unauffälliger Testzeit bleiben beispielsweise Probleme bei den einzelnen motorischen Aufgaben unentdeckt. Daher wird die bisher als eher zusätzlich erachtete Beurteilung des Ausführungserfolges bzw. -qualität der einzelnen Items eindeutig empfohlen (s. Tab. 14, S. 66).

Die Befunde zeigen, dass wahrscheinlich aufgrund der im Vergleich zum TUG höheren Kraftanteile Männer den Parcours schneller durchlaufen als Frauen (s. Tab. 28, S. 96). Dieser geschlechtsspezifische Aspekt sollte bei zukünftigen Testmodifikationen Berücksichtigung finden.

Der in der Untersuchung zur Konstruktvalidität gefundene Faktor wurde anhand der Zeiten bestimmt, die die Testpersonen für jedes einzelne Item im Durchschnitt benötigten. Hierbei sind die Ergebnisse der Faktorenanalysen, insbesondere Faktorladungen und Modelstruktur, in hohem Maße von den gesetzten Start- und Endpunkten der Zeitmessung abhängig (s. Tab. 32, S. 110). Diese müssen in weiteren Untersuchungen bzw. Berechnungen dementsprechend exakt und standardisiert umgesetzt werden.

Die Cut-Off-Werte wurden im Hinblick auf drei Fragestellungen bestimmt. Aufgrund der Befundlage (s. Abb. 29 - 34, S. 100 ff.) sollten zukünftig bei der Grenzwertbestimmung zusätzlich altersspezifisch differenzierte Cut-Off-Werte in Betracht bezogen werden. Da zudem die Cut-Off-Werte Gruppen unabhängig berechnet wurden, können bestimmte Personen- bzw. Patientengruppen isoliert nicht davon profitieren. Daher erscheint bei spezifischen Gruppen die Bestimmung von differenzierteren Cut-Off-Werten sinnvoll. Zur Verwendung des Referenzstandards „Gesundheitsstatus“ war eine Dichotomisierung des Merkmals in zwei extreme Gegenpole (gesund, krank) erforderlich (s. Kap. 5.7.2.2.3, S. 142), die nicht immer klar zu definieren sind. Nach dem Salutogenese-Modell nach Antonovsky (1979) ist der Gesundheitszu-

stand einer Person jedoch nur schlecht zu dichotomisieren. Demnach kann eine „gesunde“ Person als nicht völlig frei von Krankheit charakterisiert werden.

Weiterhin können die vermeintlich hohen prädiktiven Werte beim Referenzstandard „Gesundheitsstatus“ (s. Tab. 58, S. 145) aus der hohen Prävalenz an Personen mit körperlichen Einschränkungen resultieren. Dies könnte in weiteren Untersuchungen durch eine Parallelisierung der Personengruppen während der Rekrutierungsphase verhindert werden.

Nach der Testdurchführung stellt sich in der Therapie die Frage, wie die erhaltenen Informationen weitergehend verwertet werden (s. Abb. 9, S. 30). Im Assessmentprozess können – analog zum TUG – Personen mit positivem Ergebnis beim TUG<sub>plus</sub> im Sinne sequentieller Diagnosestrategien weiter getestet werden (American Geriatric Society, British Geriatric Society, American Academy of Orthopaedic Surgeons Panel on Falls Prevention, 2001; Bös, Wydra, & Karisch, 1992, S. 110).

## 7 Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurde der in der Praxis etablierte klassische Timed Up and Go Test (TUG) durch Integration mehrerer alltagsorientierter Bewegungshandlungen weiterentwickelt (Timed Up and Go Test plus – TUG<sub>plus</sub>). Nach einer Vor- bzw. Paralleluntersuchung zum TUG zur Überprüfung der in der Literatur festgelegten Normwerte wurden Lösungshäufigkeiten, Testzeiten sowie Testgüte überprüft und Grenzwerte anhand von Ergebnisverteilungen und der diagnostischen Güte bestimmt.

### **Vor- bzw. Paralleluntersuchung Timed Up and Go Test**

In der Untersuchung wurden durchschnittliche Testzeiten beim klassischen TUG für verschiedene Altersgruppen und Personengruppen (Patienten und „Körperlich Gesunde“) überprüft. Hierzu absolvierten 1023 Personen (♀ 54 %; Alter  $56 \pm 11$  Jahre) den Test.

Zur Absolvierung des Tests benötigten Personen jünger als 62 Jahre insgesamt durchschnittlich weniger Zeit als Personen über 62 Jahre. In Abhängigkeit der Personengruppen konnte lediglich bei Patienten mit HNO- und inneren Erkrankungen eine mit zunehmendem Alter geringfügig ansteigende Testzeit nachgewiesen werden.

„Körperlich Gesunde“ durchliefen den TUG am schnellsten (5,5 s). Sie waren schneller als Patienten mit HNO- (6,2 s), neurologischen (6,9 s), inneren (6,4 s) und orthopädischen (6,6 s) Erkrankungen ( $p \leq 0,001$ ). Neurologische Patienten waren langsamer als Patienten mit inneren bzw. HNO-Erkrankungen ( $p \leq 0,001$ ). Es ergaben sich keine geschlechtsabhängigen Unterschiede. Der angepasste Normwertbereich für die Altersspanne 18 bis 87 Jahren liegt zwischen 6,2 und 6,7 Sekunden.

Der in den 1990er Jahren aufgestellte Grenzwert von 10 Sekunden wurde also von allen Personen deutlich unterschritten.

### **Evaluation des TUG<sub>plus</sub>**

#### Verschiedene Alters- und Personengruppen

Zuerst wurden die Lösungshäufigkeiten der Items des TUG<sub>plus</sub> überprüft und die durchschnittlich benötigten Zeiten für den Parcours erfasst.

Insgesamt durchliefen 764 Personen den Parcours (Neurologie:  $n = 169$ ; Hals-Nasen-Ohrenheilkunde:  $n = 224$ ; Innere Erkrankungen:

n = 128; Orthopädie: n = 79; „Körperlich Gesunde“: n = 153; Sportsudierende: n = 11).

Das Screening zeigte bei den Personengruppen eine geringe Testschwierigkeit (88 % Lösungshäufigkeit). 94 Personen mit höhergradigen körperlichen Einschränkungen bzw. Erkrankungen konnten nicht alle Items absolvieren. Aufgrund von Multimorbidität bleibt jedoch unklar, welche genauen Ursachen zu Problemen bei bestimmten Items bzw. zur Nichtabsolvierung führten.

Weiterhin zeigte sich, dass eine ökonomische Ausführung der Items mit einer schnellen Testzeit korreliert. Hinsichtlich der Testzeit durchliefen Personen, die alle Items absolvierten, den Parcours in durchschnittlich 22 s. Insgesamt zeigte sich, dass „Körperlich Gesunde“ (18 s) schneller waren als Patienten mit neurologischen (26 s), inneren (24 s) und orthopädischen Erkrankungen (26 s). Patienten mit inneren und orthopädischen Erkrankungen waren langsamer als HNO-Patienten (21 s).

Frauen (24 s) benötigten durchschnittlich mehr Zeit als Männer (20 s). Neurologie- und Orthopädie-Patientinnen zeigten die höchsten Testzeiten (29 s). Weiterhin ist außer bei orthopädischen Patienten eine mit zunehmendem Alter ansteigende Testzeit festzustellen.

Der Normwertbereich für Personen im Alter zwischen 20 und 85 Jahren liegt zwischen 18,4 und 26,2 Sekunden.

### Testgüte

*Interne Konsistenz:* Von 166 Personen (Neurologie: n = 45; HNO: n = 59; Innere Medizin: n = 25; Orthopädie: n = 26; Körperlich Gesunde: n = 11) wurden die benötigten Zeiten für jedes einzelne Item gemessen. Alle Items zeigten mittlere Korrelationen zueinander. Mit einem  $\alpha = 0,85$  kann die interne Konsistenz als gut beurteilt werden. Die Trennschärfekoeffizienten der Items ( $r_{it} = 0,64 - 0,79$ ) bestätigen die Homogenität des Verfahrens. Durch Weglassen des Items „Aufstehen (Bank), Gehen“ könnte die Trennschärfe geringfügig verbessert werden, was zusätzlich die Testökonomie erhöhen würde.

*Paralleltest-Reliabilität:* 49 Probanden wurden randomisiert in zwei Gruppen eingeteilt. Beide Gruppen absolvierten den TUG<sub>plus</sub> in unterschiedlicher Itemreihenfolge jeweils zweimal hintereinander. Bei allen Probanden wurde bei jedem Durchgang die benötigte Testzeit notiert. Zusätzlich wurde der subjektive Anstrengungsgrad der Items anhand der Borg-Skala (6 - 20) erhoben. Es bestanden hohe Zusammenhänge ( $r = 0,87$ ;  $p \leq 0,001$ ) zwischen den Zeiten der klassischen und der um-

gekehrten Parcours-Reihenfolge. Ebenfalls hohe Zusammenhänge ( $r = 0,86$ ;  $p \leq 0,001$ ) konnten hinsichtlich des subjektiven Anstrengungsempfindens gemessen werden. Die Befunde sind unabhängig von den Kovariaten Alter, BMI und Gesundheitsstatus. Somit sind die Ergebnisse des TUG<sub>plus</sub> unabhängig davon, in welcher Reihenfolge der Parcours durchlaufen wird.

*Responsivität:* Anhand von 126 Patienten (49 % Neurologie, 52 % HNO) wurde überprüft, ob die Testergebnisse durch vorheriges Üben beeinflusst werden (Testeffekt). Zudem sollte die Frage beantwortet werden, ob anhand des TUG<sub>plus</sub> Änderungen der Motorik nach zweiwöchiger Trainingsintervention aufgezeigt werden können.

Aufgrund hoher Zusammenhänge können Testeffekte weitestgehend ausgeschlossen werden. Alle Gruppen verbesserten sich durchschnittlich nach der Trainingsintervention ( $p \leq 0,05$ ). HNO-Patienten (7 %) verbesserten sich weniger als NE-Patienten (15 %). Welche Therapieformen zu den Leistungssteigerungen führten, konnte nicht beantwortet werden.

*Kriteriumsvalidität:* Die Kriteriumsvalidität wurde anhand folgender Kriterien beurteilt: Fahrradergometrie (Stufentest: W/kg KG;  $n = 37$ ;  $57 \pm 14$  Jahre), Timed Up and Go Test (Testzeit in Sekunden;  $n = 520$ ;  $56 \pm 11$  Jahre), Motorische Basisdiagnostik (Summenscores Ausdauer, Kraft, Beweglichkeit, Koordination, Gesamtsummenscore;  $n = 119$ ;  $52 \pm 9$  Jahre), Mobilitäts-Funktionalitäts-Index (Punktescore;  $n = 226$ ;  $55 \pm 11$  Jahre), Handkraft (rechte und linke Hand: kg;  $n = 74$ ;  $52 \pm 10$  Jahre).

Der TUG<sub>plus</sub> zeigte zu allen Kriterien mittlere Korrelationen ( $p \leq 0,05$ ). Der größte Zusammenhang findet sich zur körperlichen Leistungsfähigkeit (W/kg KG).

*Konstruktvalidität:* In dieser Untersuchung sollte das dem TUG<sub>plus</sub> zugrunde liegende inhaltliche Konstrukt faktorenanalytisch untersucht werden. Aufgrund der ICF-Orientierung sollte überprüft werden, ob die Items des TUG<sub>plus</sub> das Konstrukt „Mobilität“ tatsächlich abbilden. Hierzu durchliefen 166 Probanden (Neurologie:  $n = 45$ ; HNO:  $n = 59$ ; Innere Medizin:  $n = 25$ ; Orthopädie:  $n = 26$ ; „Körperlich Gesunde“:  $n = 11$ ) den Parcours. Nach Bestätigung der Eignung der Items wurden explorative (EFA) und konfirmatorische Faktorenanalysen (KFA) für das Gesamtmodell („Körperlich Gesunde“ und „Kranke“) und jeweils für „Körperlich Kranke“ und „Körperlich Gesunde“ durchgeführt. In allen drei Modellen ergaben sich hohe Reliabilitäten ( $\alpha = 0,80 - 0,85$ ). Die EFA lieferte eine einfaktorielle Struktur. Dieser Faktor erklärt 67 - 71 % der Modelle. Die

Items laden zwischen  $\lambda = 0,68 - 0,88$  auf diesem Faktor. Bei der konfirmatorischen Betrachtung zeigten die drei Modelle anhand des  $\chi^2$ -Wertes gute Modellanpassungen. Die Anpassung anhand der berechneten Fit Indizes kann als gut charakterisiert und der gefundene Faktor demnach inhaltlich als „Mobilität“ bezeichnet werden.

## **Grenzwertbestimmung**

### Ergebnisverteilung

Die Verteilungen der Testzeiten der Gesamtstichprobe ( $n = 670$ ) dienten zur Bestimmung eines ersten Grenzwertes. Anhand des oberen Quartils findet sich ein Grenzwert von 27 Sekunden. Dieser ist unabhängig von personbezogenen Merkmalen.

### Diagnostische Güte

Die diagnostische Güte wurde anhand der Referenzstandards „Gesundheitsstatus“ und „Motorische Leistungsfähigkeit“ beurteilt, die zunächst getrennt und anschließend als kombinierte Einschränkungen betrachtet wurden.

### *Referenz „Gesundheitsstatus“*

Es sollte die Frage beantwortet werden, ob anhand des TUG<sub>plus</sub> Personen mit von Personen ohne körperliche Einschränkungen unterschieden werden können. Hierzu absolvierten 221 Personen ohne und 363 Personen mit körperlichen Beeinträchtigungen (Neurologie:  $n = 152$ ; HNO:  $n = 85$ ; Innere Medizin:  $n = 90$ ; Orthopädie:  $n = 36$ ) den Parcours.

Personen ohne körperliche Einschränkungen benötigten durchschnittlich weniger Zeit beim TUG<sub>plus</sub> ( $p \leq 0,001$ ; 18 s) als Personen mit körperlichen Einschränkungen (25 s). Anhand des Youden-Index resultiert ein Cut-Off-Wert von 21 Sekunden (69 % Sensitivität; 72 % Spezifität; AUC = 0.67; Youden-Index = 0.41; Positive/negative Likelihood Ratio = 2,5/ 0,4).

Ab einer Testzeit von 21 Sekunden zeigten sich ebenfalls höhere Testzeiten beim TUG. So benötigten Personen, die den TUG<sub>plus</sub> schneller als 21 Sekunden absolvierten weniger Zeit ( $p \leq 0,001$ ) für den TUG ( $n = 264$ ;  $5,3 \pm 1,0$  s) als Personen, die langsamer als 21 Sekunden waren ( $n = 287$ ;  $6,8 \pm 1,4$  s).

Weiterhin konnten Personen, die weniger als 21 Sekunden benötigten ( $n = 55$ ) durchschnittlich mehr Items der Motorischen Basisdiagnostik

lösen (83 %) als Personen, die langsamer waren ( $n = 64$ ; 60 %). Der größte Unterschied zeigte sich hinsichtlich des Items „Laufen mit 7 km/h“.

#### *Referenz „Motorische Leistungsfähigkeit“:*

Die Fähigkeit bei einer stufenförmigen Belastung eine Geschwindigkeit von 7 km/h zwei Minuten lang zu bewältigen, diene als weiterer Referenzstandard zur Bestimmung eines Cut-Off-Wertes (Neurologie:  $n = 67$ ; HNO:  $n = 52$ ).

53 % der Probanden schafften diese Geschwindigkeit und benötigten für den TUG<sub>plus</sub> durchschnittlich 20 Sekunden. Anhand des Youden-Index kann ein Cut-Off-Wert von 25 Sekunden identifiziert werden (Sensitivität 68 %, Spezifität 81 %, AUC = 0,73, Ratz-Index = 0,55).

#### *Referenz „Motorische Fertigkeit“ und „Gesundheitsstatus“*

Abschließend wurden beide Referenzstandards zusammengefasst (negatives Ergebnis: Fehlen körperlicher Beeinträchtigung und 7 km/h geschafft; positives Ergebnis: Vorhandensein körperlicher Beeinträchtigung und 7 km/h nicht geschafft).

Es konnte ein Cut-Off-Wert bei 21 Sekunden berechnet werden (88 % Sensitivität und Spezifität; LR+ = 7,3; LR- = 0,1; Ratz-Index = 0,82). Bei dieser Zeit erzielten 20 Personen ein negatives ( $< 21$  s) und 45 Personen ein positives Testergebnis ( $> 21$  s).

Es zeigte sich, dass die Bestimmung eines Grenzwertes, insbesondere bei der diagnostischen Güte, in hohem Maße von der Wahl des Referenzstandards abhängig ist. Insgesamt kann festgehalten werden, dass die Zusammenführung von zwei Referenzstandards ein personenbezogenes Merkmal verschärfen kann und somit zu einer Verbesserung der Entdeckungswahrscheinlichkeiten führt.

Zusammenfassend können folgende Schlussfolgerungen gezogen werden:

- Der klassische TUG hat sich in der Praxis bewährt, da dieser nur den Messparameter Zeit beinhaltet und dadurch sehr ökonomisch durchführbar ist. Der Test erscheint auch unter Verwendung des angepassten Normbereiches eher für weniger belastbare bzw. hochaltrige Personen geeignet.



- Der im Rahmen dieser Arbeit entwickelte TUG<sub>plus</sub> stellt eine Erweiterung des TUG dar. Das neue Screening beinhaltet vermehrt alltagsorientierte Bewegungsaufgaben und ermöglicht motorisch differenziertere Aussagen, erfordert allerdings einen etwas höheren Aufwand. Neben dem Faktor Zeit erfolgt eine einfache Beurteilung der Bewegungsqualität, was die Erfassung grober motorischer Auffälligkeiten ermöglicht.
- Das Screening erfüllt die Testgütekriterien Reliabilität und Validität. Für die Praxis wird ein Cut-Off-Wert von 21 Sekunden empfohlen, um Auffälligkeiten der Mobilität identifizieren zu können. Je nach Fragestellung muss der Cut-Off-Wert angepasst werden (z. B. motorische Leistungsfähigkeit 25 s).

Beim TUG<sub>plus</sub> konnten aufgrund höherer Kraftkomponenten im Vergleich zum TUG geschlechts- und personengruppen spezifische Aspekte herausgearbeitet werden. Als Konsequenz erfordert dies die weitere Bestimmung differenzierter Cut-Off-Werte (z. B. geschlechts-, alters-, krankheitsspezifisch).

In diesem Rahmen bleibt ebenfalls offen, welche körperlichen und motorischen Einschränkungen tatsächlich zu höheren Testzeiten beim klassischen TUG und beim TUG<sub>plus</sub> führen.

Weiterhin sollte die Auswirkung unterschiedlicher Testinstruktionen auf Testzeit und Bewegungsausführung bei beiden Screeningtests überprüft werden.

Die im Rahmen dieser Arbeit entwickelte Testmodifikation liefert die Basis für Weiterentwicklungen z. B. bei verschiedenen Altersklassen, insbesondere Kinder, Jugendliche und Hochaltrige sowie für spezielle Krankheitsbilder.

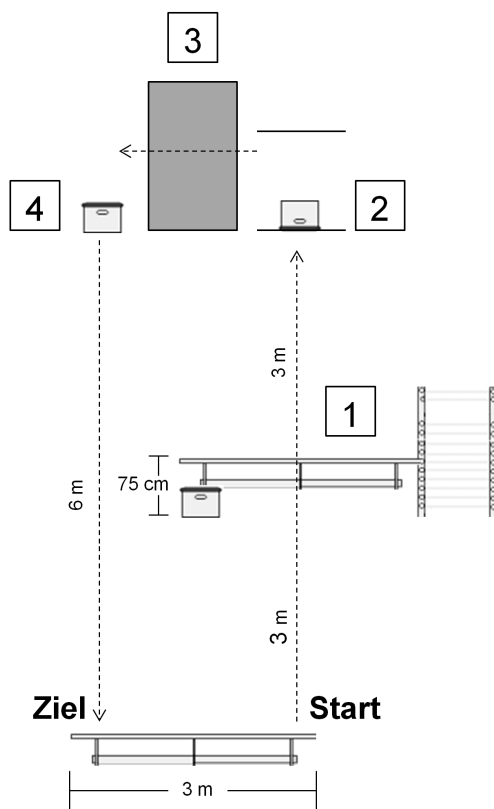
Im Folgenden werden als praktisches Fazit dieser Arbeit eine Testanleitung und ein -protokoll zum Timed up and go Test plus (TUG<sub>plus</sub>) für Erwachsene aufgeführt.

## Timed Up and Go Test plus (TUG<sub>plus</sub>)

### Screening zur Beurteilung der Mobilität bei Erwachsenen

#### – Testmanual –

**Materialbedarf:** Zwei Turnbänke, eine Sprossenwand, drei kleine Turnkästen, eine kleine Turnmatte, Stoppuhr



**Bewegungsaufgaben (Items):**

Ausgangsposition: Auf der Bank sitzend

- (1) Hindernis überwinden
- (2) Kasten umstoßen
- (3) Hinsetzen, Aufstehen (Boden)
- (4) Kasten (18 kg) tragen

*Alternativen:*

- a) 18 kg schwerer Turnkasten oder 12 kg und 6 kg schwere Getränkekiste tragen
- b) Wegstrecke wird zweimal bewältigt: 12 kg Getränkekiste 6 m tragen, zurückgehen und die restlichen 6 kg tragen
- c) Die Wegstrecke wird dreimal bewältigt: jeweils eine 6 kg schwere Getränkekiste tragen.

Endposition: Auf die Bank setzen

#### Verwendete ICF-Kategorien:

- d410 Eine elementare Körperposition ändern (Sitzen; Stehen)
- d430 Gegenstände anheben und tragen (Anheben; mit den Händen tragen; Gegenstände absetzen)
- d435 Gegenstände mit den unteren Extremitäten bewegen (Stoßen)
- d450 Gehen (Kurze Entfernungen gehen)
- d455 Sich auf andere Weise fortbewegen (Klettern/Steigen)

**Vorversuch:** Erproben Sie die einzelnen Items – wie lassen sich die Bewegungsaufgaben am besten bewältigen?

**Testanweisung:** Absolvieren Sie die Items möglichst vollständig und versuchen Sie dabei so schnell wie möglich voran zu kommen.

**Zeitmessung:** Der Teilnehmer sitzt auf der Bank. Auf das Kommando „Auf die Plätze, fertig, los“ startet die Zeitmessung. Sobald die Testperson nach Absolvierung aller Items wieder auf der Bank sitzt, wird die Zeit gestoppt.

## – Testprotokoll –

Alle Items absolviert: ☐ Nicht absolvierte Items: \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

**1. Testzeit** (alle Items absolviert)

Testzeit (s)	Bewertung
_____ s	$\leq 21$ s <input type="checkbox"/> <b>Uneingeschränkte Mobilität</b>
	$> 21$ s <input type="checkbox"/> <b>Eingeschränkte Mobilität</b>
<b>Bemerkungen:</b>	

**2. Bewegungsqualität**

Item	Bewertungsschlüssel (Punkte)	Bewertung (Punkte)
<b>1) Hindernis überwinden</b>	<b>0</b> = alleine nicht möglich <b>1</b> = Bein wir mit Hand über die Bank gezogen <b>2</b> = Bank sitzend übersteigen <b>3</b> = Bank seitlich überhocken	
<b>2) Kasten umstoßen</b>	<b>0</b> = alleine nicht möglich <b>1</b> = Mehr als 2 Kastenberührungen <b>2</b> = 2 Kastenberührungen <b>3</b> = 1 Kastenberührung	
<b>3) Hinsetzen, Aufstehen (Boden)</b>	<b>0</b> = alleine nicht möglich <b>1</b> = Proband landet auf Rücken <b>2</b> = ohne Probleme gelöst	
<b>4) Kasten tragen</b>	<b>0</b> = nicht möglich <b>1</b> = 3 x 6 kg <b>2</b> = 2 x 6 kg + 6 kg <b>3</b> = 18 kg oder 12 + 6 kg	

## Literaturverzeichnis

- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., ... & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist*, 34, 341 - 382.
- Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, J. D., Montoye, H. J., Sallis, J. F., & Paffenbarger, J. R. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*, 25, 71 - 80.
- Akobeng, A. K. (2007). Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatrica*, 96, 338 - 341.
- Alexandre, T. S., Meira, D. M., Rico, N. C., & Mizuta, S. K. (2012). Accuracy of Timed Up and Go Test for screening risk of falls among community-dwelling elderly. *Brazilian Journal of Physical Therapy*, 16, 381 - 388.
- Almeida, G. J., Schroeder, C. A., Gil, A. B., Fitzgerald, G. K., & Piva, S. R. (2010). Interrater reliability and validity of the stair ascend/descend test in subjects with total knee arthroplasty. *Archives of Physical Medicine and Rehabilitation*, 91, 932 - 938.
- Al Snih, S., Markides, K. S., Ottenbacher, K. J., & Raji, M. A. (2004). Hand grip strength and incident ADL disability in elderly Mexican Americans over a seven-year period. *Aging Clinical and Experimental Research*, 16, 481 - 486.
- Alzayer, L., Beninato, M., & Portney, L. G. (2009). The accuracy of individual Berg Balance Scale items compared with the total Berg score for classifying people with chronic stroke according to fall history. *Journal of Neurologic Physical Therapy*, 33, 136 - 143.
- American Geriatric Society, British Geriatric Society, & American Academy of Orthopaedic Surgeons Panel on Falls Prevention (2001). Guideline for the prevention of falls in older persons. *Journal of the American Geriatrics Society*, 49, 664 - 672.
- Antonovsky, A. (1979). *Health, stress, and coping*. San Francisco: Jossey-Bass.
- Arnold, C. M., & Faulkner, R. A. (2007). The history of falls and the association of the timed up and go test to falls and near-falls in older adults with hip osteoarthritis. *BMC Geriatrics*, 7, 1 - 9.

- Ashford, S., & De Souza, L. (2000). A comparison of the timing of muscle activity during sitting down compared to standing up. *Physiotherapy Research International*, 5, 111 - 128.
- Asmidawati, A., Hamid, T. A., Hussain, R. M., & Hill, K. D. (2014). Home based exercise to improve turning and mobility performance among community dwelling older adults: protocol for a randomized controlled trial. *BMC Geriatrics*, 14, 1 - 9.
- Baker, D. W. (2006). The meaning and the measure of health literacy. *Journal of General Internal Medicine*, 21, 878 - 883.
- Baldus, A., Schüle, K., & Huber, G. (2012). ICF-orientierte Qualitätssystematik in der Rehabilitation. In K. Schüle, & G. Huber (Hrsg.), *Grundlagen der Sport- und Bewegungstherapie. Prävention, ambulante und stationäre Rehabilitation* (S. 325 - 329). Köln: Deutscher Ärzte-Verlag.
- Ballreich, R. (1970). *Grundlagen sportmotorische Tests*. Frankfurt: Limpert.
- Balogun, J. A., Akomolafe, C. T., & Amusa, L. O. (1991). Grip strength: effects of testing posture and elbow position. *Archives of Physical Medicine and Rehabilitation*, 72, 280 - 283.
- Barry, E., Galvin, R., Keogh, C., Horgan, F., & Fahey, T. (2014). Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. *BMC Geriatrics*, 14, 1 - 14.
- Bennell, K., Dobson, F., & Hinman, R. (2011). Measures of physical performance assessments: Self-Paced Walk Test (SPWT), Stair Climb Test (SCT), Six-Minute Walk Test (6MWT), Chair Stand Test (CST), Timed Up & Go (TUG), Sock Test, Lift and Carry Test (LCT), and Car Task. *Arthritis Care & Research*, 63, 350 - 370.
- Bennie, S., Bruner, K., Dizon, A., Fritz, H., Goodman, B., & Peterson, S. (2003). Measurements of balance: comparison of the "Timed Up and Go" test and Functional Reach test with the Berg Balance Scale. *Journal of Physical Therapy Science*, 15, 93 - 97.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588 - 606.
- Berg, K., Wood-Dauphinee, S. L., Williams, J. I., & Gayton, D. (1989). Measuring balance in the elderly: preliminary development of an instrument. *Physiotherapy Canada*, 41, 304 - 311.

- Berg, K. O., Wood-Dauphinee, S. L., & Williams, J. I. (1995). The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scandinavian Journal of Rehabilitation Medicine*, 27, 27 - 36.
- Berg, W. A., Gutierrez, L., Ness Aiver, M. S., Carter, W. B., Bhargavan, M., Lewis, R. S., & Ioffe, O. B. (2004). Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in pre-operative assessment of breast cancer. *Radiology*, 233, 830 - 849.
- Bird, M. L., Hill, K. D., & Fell, J. W. (2012). A randomized controlled study investigating static and dynamic balance in older adults after training with pilates. *Archives of Physical Medicine and Rehabilitation*, 93, 43 - 49.
- Bischoff, H. A., Stähelin, H. B., Monsch, A. U., Iversen, M. D., Weyh, A., Dechend, M., ... & Theiler, R. (2003). Identifying a cut-off point for normal mobility: A comparison of the timed 'up and go' test in community-dwelling and institutionalised elderly women. *Age and Ageing*, 32, 315 - 320.
- Blennerhassett, J., & Dite, W. (2004). Additional task-related practice improves mobility and upper limb function early after stroke: a randomised controlled trial. *Australian Journal of Physiotherapy*, 50, 219 - 224.
- Bös, K., & Mechling, H. (1983). *Dimensionen sportmotorischer Leistungen*. Schorndorf: Hofmann.
- Bös, K. (1987). *Handbuch sportmotorischer Tests*. Göttingen: Hogrefe.
- Bös, K., Wydra, G., & Karisch, G. (1992). *Gesundheitsförderung durch Bewegung, Spiel und Sport. Ziele und Methoden des Gesundheits-sports in der Klinik*. Erlangen: Perimed.
- Bös, K., Hänsel, F., & Schott, N. (2004). *Empirische Untersuchungen in der Sportwissenschaft: Planung - Auswertung - Statistik* (2. Aufl.). Hamburg: Czwalina.
- Bös, K. (2017). *Handbuch Motorische Tests. Sportmotorische Tests, Motorische Funktionstests, Fragebögen zur körperlich-sportlichen Aktivität und sportpsychologische Diagnoseverfahren* (3. Aufl.). Göttingen: Hogrefe.
- Bohannon, R. W. (2006). Reference values for the timed Up and Go test: a descriptive meta-analysis. *Journal of Geriatric Physical Therapy*, 29, 64 - 68.
- Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise*, 14, 377 - 381.

- Bortz, J., & Döring, N. (2005). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... & de Vet, H. C. W. (2003). Standards for reporting of diagnostic accuracy (STARD) group. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Annals of Internal Medicine*, 138, 40 - 44.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A., & Cohen, J. F. (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *Radiology*, 277, 826 - 832.
- Botolfsen, P., Helbostad, J. L., Moe-Nilssen, R., & Wall, J. C. (2008). Reliability and concurrent validity of the Expanded Timed Up-and-Go test in older people with impaired mobility. *Physics Research International*, 13, 94 - 106.
- Boutin, A., Panzer, S., & Blandin, Y. (2013). Retrieval practice in motor learning. *Human Movement Science*, 32, 1201 - 1213.
- Brisswalter, J., & Mottet, D. (1996). Energy cost and stride duration variability at preferred transition gait speed between walking and running. *Canadian Journal of Applied Physiology*, 21, 471 - 480.
- Broekmans, T., Roelants, M., Alders, G., Feys, P., Thijs, H., & Eijnde, B. O. (2010). Exploring the effects of a 20-week whole-body vibration training program on leg muscle performance and function in persons with multiple sclerosis. *Journal of Rehabilitation Medicine*, 42, 866 - 872.
- Burton, A. W., & Miller, D. E. (1998). *Movement skill assessment*. Champaign: Human Kinetics.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Butland, R. J., Pang, J. A. C. K., Gross, E. R., Woodcock, A. A., & Geddes, D. M. (1982). Two-, six-, and 12-minute walking tests in respiratory disease. *British Medical Journal (Clinical Research Ed.)*, 284, 1607.
- Cadore, E. L., Casas-Herrero, A., Zambom-Ferraresi, F., Idoate, F., Millor, N., Gómez, M., ... & Izquierdo, M. (2014). Multicomponent exercises including muscle power training enhance muscle mass, power output, and functional outcomes in institutionalized frail nonagenarians. *Age*, 36, 773 - 785.

- Canbek, J., Fulk, G., Nof, L., & Echternach, J. (2013). Test-retest reliability and construct validity of the tinetti performance-oriented mobility assessment in people with stroke. *Journal of Neurologic Physical Therapy*, 37, 14 - 19.
- Cesari, M., Kritchevsky, S. B., Penninx, B. W., Nicklas, B. J., Simonsick, E. M., Newman, A. B., ... & Visser, M. (2005). Prognostic value of usual gait speed in well-functioning older people - results from the health, aging and body composition study. *Journal of the American Geriatrics Society*, 53, 1675 - 1680.
- Chiu, A. Y. Y., Au-Yeung, S. S. Y., & Lo, S. K. (2003). A comparison of four functional tests in discriminating fallers from non-fallers in older people. *Disability and Rehabilitation*, 25, 45 - 50.
- Chou, C. Y., Chien, C. W., Hsueh, I. P., Sheu, C. F., Wang, C. H., & Hsieh, C. L. (2006). Developing a short form of the Berg Balance Scale for people with stroke. *Physical Therapy*, 86, 195 - 204.
- Clark, W. A., Van Ham, M., & Coulter, R. (2014). Spatial mobility and social outcomes. *Journal of Housing and the Built Environment*, 29, 699 - 727.
- Clauß, G., & Ebner, H. (1975). *Grundlagen der Statistik*. Frankfurt am Main: Harri Deutsch.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Erlbaum.
- Collin, C., Wade, D. T., & Cochrane, G. M. (1992). Functional outcome of lower limb amputees with peripheral vascular disease. *Clinical Rehabilitation*, 6, 13 - 21.
- Collins, E., O'Connell, S., Jelinek, C., Miskevics, S., & Budiman-Mak, E. (2008). Evaluation of psychometric properties of Walking Impairment Questionnaire in overweight patients with osteoarthritis of knee. *Journal of Rehabilitation Research and Development*, 45, 559 - 565.
- Commission on Chronic Illness (1957). *Chronic illness in the United States: Volume I. Prevention of chronic illness*. Cambridge: Harvard University Press.
- Corrie, H., Brooke-Wavell, K., Mansfield, N. J., Cowley, A., Morris, R., & Masud, T. (2014). Effects of vertical and side-alternating vibration training on fall risk factors and bone turnover in older people at risk of falls. *Age and Ageing*, 44, 115 - 122.
- Croyle, R. T. (1995). *Psychosocial effects of screening for disease prevention and detection*. Oxford: University Press.
- Cumming, R. G., Salkeld, G., Thomas, M., & Szonyi, G. (2000). Prospective study of the impact of fear of falling on activities of daily living, SF-36



- scores, and nursing home admission. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 55, M299 - M305.
- Dallmeijer, A. J., & Van der Woude, L. H. V. (2001). Health related functional status in men with spinal cord injury: relationship with lesion level and endurance capacity. *Spinal Cord*, 39, 577 - 583.
- Datta, D., Ariyaratnam, R., & Hilton, S. (1996). Timed walking test – an all-embracing outcome measure for lower-limb amputees? *Clinical Rehabilitation*, 10, 227 - 232.
- Dean, C. M., Richards, C. L., & Malouin, F. (2000). Task-related circuit training improves performance of locomotor tasks in chronic stroke: A randomized, controlled pilot trial. *Archives of Physical Medicine and Rehabilitation*, 81, 409 - 417.
- De Araújo, T. B., de Oliveira, R. J., Martins, W. R., de Moura Pereira, M., Copetti, F., & Safons, M. P. (2013). Effects of hippotherapy on mobility, strength and balance in elderly. *Archives of Gerontology and Geriatrics*, 56, 478 - 481.
- Deathe, A. B., & Miller, W. C. (2005). The L test of functional mobility: measurement properties of a modified version of the timed “up & go” test designed for people with lower-limb amputations. *Physical Therapy*, 85, 626 - 635.
- De Marées, H. (2003). *Sportphysiologie*. Köln: Sportverlag Strauß.
- Demers, C., McKelvie, R. S., Negassa, A., & Yusuf, S. (2001). Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. *American Heart Journal*, 142, 698 - 703.
- Denegar, C. R., & Fraser, M. (2006). How useful are physical examination procedures? Understanding and applying likelihood ratios. *Journal of Athletic Training*, 41, 201 - 206.
- Deutsches Institut für Medizinische Dokumentation und Information (DIMDI) (2005). *ICF – Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit*. Köln: DIMDI.
- Deutsche Rentenversicherung Bund (DRV) (2015a). *Sozialgesetzbuch*. Berlin: DRV.
- Deutsche Rentenversicherung (DRV) (2015b). *Die medizinische und berufliche Rehabilitation der Rentenversicherung im Licht der Statistik*. Berlin: DRV.
- Deutsches Zentrum für Altersfragen (2014). *Deutscher Alterssurvey 2014. Zentrale Befunde*. Berlin: Deutsches Zentrum für Altersfragen.

- Dimeo, F., Stieglitz, R. D., Novelli-Fischer, U., Fetscher, S., Mertelsmann, R., & Keul, J. (1997). Correlation between physical performance and fatigue in cancer patients. *Annals of Oncology*, 8, 1251 - 1255.
- Doyle, P. J. (2002). Measuring health outcomes in stroke survivors. *Archives of Physical Medicine and Rehabilitation*, 83, S39 - S43.
- Dunlop, D. D., Manheim, L. M., Sohn, M. W., Liu, X., & Chang, R. W. (2002). Incidence of functional limitation in older adults: the impact of gender, race, and chronic conditions. *Archives of Physical Medicine and Rehabilitation*, 83, 964 - 971.
- Eekhof, J. A., De Bock, G. H., Schaapveld, K., & Springer, M. P. (2001). Functional mobility assessment at home. Timed up and go test using three different chairs. *Canadian Family Physician*, 47, 1205 - 1207.
- Evans, G. F., Anthony, T., Appelbaum, A. H., Schumpert, T. D., Levy, K. R., Amirkhan, R. H., Cambell, T. J., Lopez, J., & Turnage, R. H. (2001). The diagnostic accuracy of mammography in the evaluation of male breast disease. *The American Journal of Surgery*, 181, 96 - 100.
- Ewert, T., Allen, D. D., Wilson, M., Üstün, B., & Stucki, G. (2010). Validation of the International Classification of Functioning, Disability and Health framework using multidimensional item response modeling. *Disability and Rehabilitation*, 32, 1397 - 1405.
- Faber, M. J., Bosscher, R. J., & Van Wieringen, P. C. (2006). Clinimetric properties of the performance-oriented mobility assessment. *Physical Therapy*, 86, 944 - 954.
- Flansbjerg, U. B., Holmbäck, A. M., Downham, D., Patten, C., & Lexell, J. (2005). Reliability of gait performance tests in men and women with hemiparesis after stroke. *Journal of Rehabilitation Medicine*, 37, 75 - 82.
- Fluss, R., Faraggi, D., & Reiser, B. (2005). Estimation of the Youden Index and its Associated Cutoff Point. *Biomedical Journal*, 47, 458 - 472.
- Fortin, P. R., Clarke, A. E., Joseph, L., Liang, M. H., Tanzer, M., Ferland, D., ... & Mahomed, N. (1999). Outcomes of total hip and knee replacement. *Arthritis & Rheumatology*, 42, 1722 - 1728.
- Frontera, W. R., Hughes, V. A., Lutz, K. J., & Evans, W. J. (1991). A cross-sectional study of muscle strength and mass in 45- to 78-year-old men and women. *Journal of Applied Physiology*, 71, 644 - 650.

- Fryback, D. G., & Thornbury, J. R. (1991). The efficacy of diagnostic imaging. *Medical Decision Making*, 11, 88 - 94.
- Fuchs, J., Kuhnert, R., & Scheidt-Nave, C. (2017). 12-Monats-Prävalenz von Arthrose in Deutschland. *Journal of Health Monitoring*, 2, 55 - 60.
- Gaßner, H., Janzen, A., Schwirtz, A., & Jansen, P. (2014). Random whole body vibration over 5 weeks leads to effects similar to placebo: a controlled study in Parkinson's disease. *Parkinson's Disease*, 1 - 9.
- Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I. M., ... & Swain, D. P. (2011). Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine & Science in Sports & Exercise*, 43, 1334 - 1359.
- Geyh, S., Cieza, A., Schouten, J., Dickson, H., Frommelt, P., Omar, Z., ... & Stucki, G. (2004). ICF Core Sets for stroke. *Journal of Rehabilitation Medicine*, 36, 135 - 141.
- Giampaoli, S., Ferrucci, L., Cecchi, F., Noce, C. L., Poce, A., Dima, F., Santaquilani, A., Vescio, M. F., & Menotti, A. (1999). Hand-grip strength predicts incident disability in non-disabled older men. *British Geriatrics Society*, 28, 283 - 288.
- Gleeson, M., Sherrington, C., & Keay, L. (2014). Exercise and physical training improve physical function in older adults with visual impairments but their effect on falls is unclear: a systematic review. *Journal of Physiotherapy*, 60, 130 - 135.
- Godges, J. J., MacRae, H., Longdon, C., Tinberg, C., & MacRae, P. (1989). The effects of two stretching procedures on hip range of motion and gait economy. *Journal of Orthopaedic & Sports Physical Therapy*, 10, 350 - 357.
- Graf, C., & Halle, M. (2007). Die Bedeutung von körperlicher Aktivität bei koronarer Herzkrankheit. *Deutsche Zeitschrift für Sportmedizin*, 58, 322 - 327.
- Granacher, U., Lacroix, A., Muehlbauer, T., Roettger, K., & Gollhofer, A. (2013). Effects of core instability strength training on trunk muscle strength, spinal mobility, dynamic balance and functional mobility in older adults. *Gerontology*, 59, 105 - 113.
- Grotkamp, S., Cibis, W., Bahemann, A., Baldus, A., Behrens, J., Nyffeler, I. D., ... & Seger, W. (2014). Bedeutung der personbezogenen Faktoren der ICF für die Nutzung in der praktischen

- Sozialmedizin und Rehabilitation. *Gesundheitswesen*, 76, 172 - 180.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological Assessment*, 12, 19 - 30.
- Guralnik, J. M., Ferrucci, L., Pieper, C. F., Leveille, S. G., Markides, K. S., Ostir, G. V., ... & Wallace, R. B. (2000). Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 55, M221 - M231.
- Guyatt, G. H., Townsend, M., Berman, L. B., & Keller, J. L. (1987). A comparison of Likert and visual analogue scales for measuring change in function. *Journal of chronic diseases*, 40, 1129 - 1133.
- Haines, T., Kuys, S. S., Morrison, G., Clarke, J., & Bew, P. (2008). Balance impairment not predictive of falls in geriatric rehabilitation wards. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 63, 523 - 528.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29 - 36.
- Harada, N., Chiu, V., Damron-Rodriguez, J., Fowler, E., Siu, A., & Reuben, D. B. (1995). Screening for balance and mobility impairment in elderly individuals living in residential care facilities. *Physical Therapy*, 75, 462 - 469.
- Hardy, S. E., Dubin, J. A., Holford, T. R., & Gill, T. M. (2005). Transitions between states of disability and independence among older persons. *American Journal of Epidemiology*, 161, 575 - 584.
- Harris, J. E., Eng, J. J., Marigold, D. S., Tokuno, C. D., & Louis, C. L. (2005). Relationship of balance and mobility to fall incidence in people with chronic stroke. *Physical Therapy*, 85, 150 - 158.
- Hartley-O'Brien, S. J. (1980). Six mobilization exercises for active range of hip flexion. *Research Quarterly for Exercise and Sport*, 51, 625 - 635.
- Harun, A., Semenov, Y. R., & Agrawal, Y. (2015). Vestibular function and activities of daily living: analysis of the 1999 to 2004 National Health and Nutrition Examination Surveys. *Gerontology and Geriatric Medicine*, 1, 1 - 8.

- Hawkins, R. C. (2005). The evidence based medicine approach to diagnostic testing: practicabilities and limitations. *Clinical Biochemist Reviews*, 26, 7 - 18.
- Herdman, S. J., Blatt, P., Schubert, M. C., & Tusa, R. J. (2000). Falls in patients with vestibular deficits. *Otology & Neurotology*, 21, 847 - 851.
- Herrmann, C., & Gerlach, E. (2014). Motorische Basiskompetenzen in der Grundschule: pädagogische Zielentscheidung und Aufgabenentwicklung. *Sportunterricht*, 63, 322 - 328.
- Hillert, A., Koch, S., & Hedlund, S. (2007). *Stressbewältigung am Arbeitsplatz. Ein stationäres Gruppenprogramm*. Göttingen: Vandenhoeck & Ruprecht.
- Hirvensalo, M., Rantanen, T., & Heikkinen, E. (2000). Mobility difficulties and physical activity as predictors of mortality and loss of independence in the community-living older population. *Journal of the American Geriatrics Society*, 48, 493 - 498.
- Hoogeboom, T. J., Dronkers, J. J., van den Ende, C. H., Oosting, E., & Van Meeteren, N. L. (2010). Preoperative therapeutic exercise in frail elderly scheduled for total hip replacement: a randomized pilot trial. *Clinical Rehabilitation*, 24, 901 - 910.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, 6, 1 - 55.
- Huang, S. L., Hsieh, C. L., Wu, R. M., Tai, C. H., Lin, C. H., & Lu, W. S. (2011). Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with Parkinson disease. *Physical Therapy*, 91, 114 - 121.
- Huber, G. (2012). Sporttherapeutisches Assessment. In K. Schüle, & G. Huber (Hrsg.), *Grundlagen der Sport- und Bewegungstherapie* (S. 182 - 194). Köln: Deutscher Ärzte-Verlag.
- Hummel, A., & Borchert, T. (2015). Entwicklung motorischer Kompetenzen schließt Förderung motorischer Fähigkeiten ein. *Sportunterricht*, 64, 138 - 144.
- Hunter, G. R., Weinsier, R. L., Gower, B. A., & Wetzstein, C. (2001). Age-related decrease in resting energy expenditure in sedentary white women: effects of regional differences in lean and fat mass. *The American Journal of Clinical Nutrition*, 73, 333 - 337.
- Hyndman, D., Ashburn, A., & Stack, E. (2002). Fall events among people with stroke living in the community: circumstances of falls and

- characteristics of fallers. *Archives of Physical Medicine and Rehabilitation*, 83, 165 - 170.
- Ibrahim, A., Singh, D. K. A., & Shahar, S. (2017). 'Timed Up and Go' test: Age, gender and cognitive impairment stratified normative values of older adults. *PloS One*, 12, 1 - 14.
- Iezzoni, L. I., McCarthy, E. P., Davis, R. B., & Siebens, H. (2001). Mobility difficulties are not only a problem of old age. *Journal of General Internal Medicine*, 16, 235 - 243.
- Israel, S. (1983). Körperliche Normbereiche in ihrem Bezug zur Gesundheitsstabilität. *Theorie und Praxis der Körperkultur*, 32, 360 - 363.
- Israel, S. (1990): Bewegungsinduzierte körperliche Idealnomen als Grundlage hoher Gesundheitsstabilität. *Theorie und Praxis der Körperkultur* 39, 9 - 15.
- Jaeschke, R., Guyatt, J. H., Sacket, D. L. (1994). User's guide to the medical literature, III: how to use an article about a diagnostic test. B: What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *The Journal of the American Medical Association*, 271, 703.
- Jaeschke, R., Guyatt, G., & Lijmer, J. (2002). *Diagnostic tests. Users' Guides to the Medical Literature – A Manual for Evidence-Based Clinical Practice*. Chicago: AMA Press.
- Janssen, I., & Ross, R. (2005). Linking age-related changes in skeletal muscle mass and composition with metabolism and disease. *Journal of Nutrition Health and Aging*, 9, 408 - 419.
- Jette, A. M., Jette, D. U., Ng, J., Plotkin, D. J., Bach, M. A., & Musculoskeletal Impairment (MSI) Study Group (1999). Are performance-based measures sufficiently reliable for use in multicenter trials? *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 54, M3 - M6.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2014a). Entwicklung eines neuen ICF-basierten Mobilitätstests: Ergebnisse einer ersten Pilotstudie. In Deutsche Rentenversicherung Bund (Hrsg.), *Deutscher Kongress für Rehabilitationsforschung: Arbeit - Gesundheit - Rehabilitation* (S. 152 - 153). Berlin: DRV.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2014b). Timed Up and Go Test für Patienten mittleren Alters. *Bewegungstherapie und Gesundheitssport*, 30, 240.

- Kaczmarek, C., Fröhlich, M., Schwarz, M. & Wydra, G. (2015). Bewertungsmaßstab eines ICF orientierten Mobilitätstests. In T. Könecke, H. Preuß & W. I. Schöllhorn (Hrsg.), *Moving Minds - Crossing Boundaries in Sport Science*. 22. Sportwissenschaftlicher Hochschultag (Band 251) (S. 352). Hamburg: Czwalina.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2016a). Mobilitätskonzept in Sporttherapie und -wissenschaft.. *Sportwissenschaft*, 46, 268 - 274.
- Kaczmarek, C., Schwarz, M., Backes A., & Wydra, G. (2016b). Testgüte und Diagnostische Güte eines neuen Mobilitätsscreenings. Dvs-Jahrestagung Bewegung, Raum und Gesundheit September 2016 in Karlsruhe.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2017a). Teststatistische Überprüfung eines neuen ICF orientierten, modifizierten Timed Up and Go Tests. In Deutsche Rentenversicherung Bund (Hrsg.), *Deutscher Kongress für Rehabilitationsforschung. Prävention und Rehabilitation in Zeiten der Globalisierung* (S. 71 - 72). Frankfurt: DRV-Schriften.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2017b). TUG 50+ – Ein neues Screening für die Bewegungstherapie. In A. Schwirtz, F. Mess, Y. Demetriou & V. Senner (Hrsg.), *Innovation & Technologie im Sport* (S. 274). Hamburg: Czwalina.
- Kaczmarek, C., Schwarz, M., & Wydra, G. (2017c). Reliabilität, Vailidität und diagnostische Güte eines neuen Mobilitätsscreenings. In H. Wäsche, G. Sudeck, R. S. Kähler, L. Vogt, & A. Woll (Hrsg.), *Bewegung, Raum und Gesundheit* (S. 143 - 152). Hamburg: Feldhaus.
- Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *The Journal of the American Medical Association*, 185, 914 - 919.
- Kegelmeyer, D. A., Kloos, A. D., Thomas, K. M., & Kostyk, S. K. (2007). Reliability and validity of the Tinetti Mobility Test for individuals with Parkinson disease. *Physical Therapy*, 87, 1369 - 1378.
- Kehayias, J. J., Fiatarone, M. A., Zhuang, H., & Roubenoff, R. (1997). Total body potassium and body fat: relevance to aging. *The American Journal of Clinical Nutrition*, 66, 904 - 910.
- Kennedy, D. M., Stratford, P. W., Wessel, J., Gollish, J. D., & Penney, D. (2005). Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskeletal Disorders*, 6, 1 - 12.

- Kindermann, W. (1987). Ergometrie-Empfehlungen für die ärztliche Praxis. *Deutsche Zeitschrift für Sportmedizin*, 38, 244 - 268.
- Klafki, W. (2005). Bewegungskompetenz als Bildungsdimension. In R. Laging, & R. Prohl (Hrsg.), *Bildung und Bewegung* (S. 15 - 24). Hamburg: Czwalina.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., ... & Vollmer, H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards – Expertise*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Klima, D. W., Newton, R. A., Keshner, E. A., & Davey, A. (2013). Fear of falling and balance ability in older men: the priest study. *Journal of Aging and Physical Activity*, 21, 375 - 386.
- Knottnerus, J. A., & Muris, J. (2002). Assessment of the accuracy of diagnostic tests: the cross-sectional study. In J. A. Knottnerus (eds.), *The evidence base of clinical diagnosis* (S. 39 - 59). London, UK: BMJ Books.
- Knottnerus, J. A., Buntinx, F., & Van Weel, C. (2009). General introduction: evaluation of diagnostic procedures. In J. A. Knottnerus, & F. Buntinx (eds.), *The evidence base of clinical diagnosis. Theory and methods of diagnostic research* (S. 1 - 19). Oxford: Wiley-Blackwell.
- Kojima, G., Masud, T., Kendrick, D., Morris, R., Gawler, S., Treml, J., & Iliffe, S. (2015). Does the timed up and go test predict future falls among British community-dwelling older people? Prospective cohort study nested within a randomised controlled trial. *BMC Geriatrics*, 15, 1 - 7.
- Kolominsky-Rabas, P. L., Sarti, C., Heuschmann, P. U., Graf, C., Siemonsen, S., Neundoerfer, B., ... & von Stockert, T. R. (1998). A prospective community-based study of stroke in Germany – the Erlangen Stroke Project (ESPro): incidence and case fatality at 1, 3, and 12 months. *Stroke*, 29, 2501 - 2506.
- Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland (KMK) (2005). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung*. München: Wolters Kluwer.
- Kristensen, M. T., Foss, N. B., & Kehlet, H. (2007). Timed “up & go” test as a predictor of falls within 6 months after hip fracture surgery. *Physical Therapy*, 87, 24 - 30.



- Kroenke, C. H., Rosner, B., Chen, W. Y., Kawachi, I., Colditz, G. A., & Holmes, M. D. (2004). Functional impact of breast cancer by age at diagnosis. *Journal of Clinical Oncology*, 22, 1849 - 1856.
- Kumar, D., Wyatt, C., Chiba, K., Lee, S., Nardo, L., Link, T. M., ... & Souza, R. B. (2015). Anatomic correlates of reduced hip extension during walking in individuals with mild-moderate radiographic hip osteoarthritis. *Journal of Orthopaedic Research*, 33, 527 - 534.
- Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia, Critical Care & Pain*, 8, 221 - 223.
- Lamb, S. E., Ferrucci, L., Volapto, S., Fried, L. P., & Guralnik, J. M. (2003). Risk factors for falling in home-dwelling older women with stroke: the Women's Health and Aging Study. *Stroke*, 34, 494 - 501.
- Lange, M. (2013). Entwicklung und Überprüfung eines Fragebogens zur Erfassung körperlicher Aktivität bei älteren Menschen (60 bis 85 Jahre). *Leipziger Sportwissenschaftliche Beiträge*, 54, 145 - 155.
- Lee, H. C., Chang, K. C., Tsauo, J. Y., Hung, J. W., Huang, Y. C., & Lin, S. I. (2013). Effects of a multifactorial fall prevention program on fall incidence and physical function in community-dwelling older adults with risk of falls. *Archives of Physical Medicine and Rehabilitation*, 94, 606 - 615.
- Leeflang, M. M., Rutjes, A. W., Reitsma, J. B., Hooft, L., & Bossuyt, P. M. (2013). Variation of a test's sensitivity and specificity with disease prevalence. *Canadian Medical Association Journal*, 185, E537 - E544.
- Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Lin, M. R., Hwang, H. F., Hu, M. H., Wu, H. D. I., Wang, Y. W., & Huang, F. C. (2004). Psychometric comparisons of the timed Up and Go, One-Leg stand, functional reach, and tinetti balance measures in community-dwelling older people. *Journal of the American Geriatrics Society*, 52, 1343 - 1348.
- Ling, C. H., Taekema, D., de Craen, A. J., Gussekloo, J., Westendorp, R. G., & Maier, A. B. (2010). Handgrip strength and mortality in the oldest old population: the Leiden 85-plus study. *Canadian Medical Association Journal*, 182, 429 - 435.
- Lipsitz, L. A., Jonsson, P. V., Kelley, M. M., & Koestner, J. S. (1991). Causes and correlates of recurrent falls in ambulatory frail elderly. *Journal of Gerontology*, 46, M114 - M122.

- Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: a review. *Psychological Bulletin*, 94, 68 - 99.
- Löllgen, H. (2010). Leistung, Anstrengungsempfinden, metabolisches Äquivalent, Lebensqualität. In H. Löllgen, E. Erdmann, & A. Gitt (Hrsg.), *Ergometrie. Belastungsuntersuchungen in Klinik und Praxis*. Heidelberg: Springer.
- Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., ... & Al Mazroa, M. A. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380, 2095 - 2128.
- Ludwig, G. (2012). Bewegungsdiagnostik. In K. Schüle, & G. Huber (Hrsg.), *Grundlagen der Sport- und Bewegungstherapie. Prävention, ambulante und stationäre Rehabilitation* (S. 173 - 181). Köln: Deutscher Ärzte-Verlag.
- Maanum, G., Jahnsen, R., FRøSLIE, K. F., Larsen, K. L., & Keller, A. (2010). Walking ability and predictors of performance on the 6-minute walk test in adults with spastic cerebral palsy. *Developmental Medicine & Child Neurology*, 52, E126 - E132.
- Mao, H. F., Hsueh, I. P., Tang, P. F., Sheu, C. F., & Hsieh, C. L. (2002). Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke*, 33, 1022 - 1027.
- Mariush, M. E. (2004). *The use of psychological testing for treatment planning and outcomes assessment*. New Jersey: Lawrence Erlbaum Associates.
- Marx, P., & Lenhard, W. (2010). Diagnostische Merkmale von Screening-Verfahren zur Früherkennung möglicher Probleme beim Schriftspracherwerb. In M. Hasselhorn, & W. Schneider (Hrsg), *Frühprognose schulischer Kompetenzen* (S. 68 - 84). Göttingen: Hogrefe.
- Mathiowetz, V., Rennells, C., & Donahoe, L. (1985). Effect of elbow position on grip and key pinch strength. *The Journal of Hand Surgery*, 10, 694 - 697.
- Mathias, S., Nayak, U. S., & Isaacs, B. (1986). Balance in elderly patients: the "get-up and go" test. *Archives of Physical Medicine Rehabilitation*, 67, 387 - 389.
- Matsuda, P. N., Shumway-Cook, A., Ciol, M. A., Bombardier, C. H., & Kartin, D. A. (2012). Understanding falls in multiple sclerosis: association of mobility status, concerns about falling, and accumulated impairments. *Physical Therapy*, 92, 407 - 415.

- McMurdo, M. E. T. & Rennie, L. M. (1993). A controlled trial of exercise by residents of old people's homes. *Age and Ageing*, 22, 11 - 15.
- Means, K. M., Rodell, D. E., & O'Sullivan, P. S. (2005). Balance, mobility, and falls among community-dwelling elderly persons: effects of a rehabilitation exercise program. *American Journal of Physical Medicine & Rehabilitation*, 84, 238 - 250.
- Merx, H., Dreinhöfer, K. E., & Günther, K. P. (2007). Sozialmedizinische Bedeutung der Arthrose in Deutschland. *Zeitschrift für Orthopädie und Unfallchirurgie*, 145, 421 - 429.
- Moore, S. C., Patel, A. V., Matthews, C. E., de Gonzalez, A. B., Park, Y., Katki, H. A., ... & Thun, M. (2012). Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. *PLoS Medicine*, 9, 1 - 14.
- Morris, W. W., Buckwalter, K. C., Cleary, T. A., Gilmer, J. S., Hatz, D. L., & Studer, M. (1990). Refinement of the Iowa self-assessment inventory. *The Gerontologist*, 30, 243 - 248.
- Morris, S., Morris, M. E., & Iansek, R. (2001). Reliability of measurements obtained with the Timed "Up & Go" test in people with parkinson disease. *Physical Therapy*, 81, 810 - 818.
- Morrison, E. J., Shoukath, S., Tansley, P., & Grinsell, D. (2013). Donor site morbidity of an islanded inferior gluteal artery myocutaneous flap with vascularized fascia lata. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 66, 962 - 967.
- Motel-Klingebiel, A., Wurm, S., & Tesch-Römer, C. (2010). *Altern im Wandel: Befunde des Deutschen Alterssurveys (DEAS)*. Kohlhammer.
- Müller-Fahrnow, W., Greitemann, B., Radoschewski, F. M., Gerwin, H., & Hansmeier, T. (2005). Berufliche Orientierung in der medizinischen Rehabilitation und Leistungen zur Teilhabe am Arbeitsleben. *Rehabilitation*, 44, E32 - E45.
- Mulasso, A., Roppolo, M., Liubicich, M. E., Settanni, M., & Rabaglietti, E. (2015). A multicomponent exercise program for older adults living in residential care facilities: direct and indirect effects on physical functioning. *Journal of Aging and Physical Activity*, 23, 409 - 416.
- Musselman, K., & Brouwer, B. (2005). Gender-related differences in physical performance among seniors. *Journal of Aging and Physical Activity*, 13, 239 - 253.

- Musterd, S., Van Gent, W. P., Das, M., & Latten, J. (2016). Adaptive behaviour in urban space: Residential mobility in response to social distance. *Urban Studies*, 53, 227 - 246.
- Ng, S. S., & Hui-Chan, C. W. (2005). The timed up & go test: its reliability and association with lower-limb impairments and locomotor capacities in people with chronic stroke. *Archives of Physical Medicine and Rehabilitation*, 86, 1641 - 1647.
- Nordenfelt, L. (2003). Action theory, disability and ICF. *Disability and Rehabilitation*, 25, 1075 - 1079.
- Nordin, E., Rosendahl, E., & Lundin-Olsson, L. (2006). Timed "Up & Go" Test: reliability in older people dependent in activities of daily living – focus on cognitive state. *Physical Therapy*, 86, 646 - 655.
- Nordin, E., Lindelöf, N., Rosendahl, E., Jensen, J., & Lundin-Olsson, L. (2008). Prognostic validity of the Timed Up-and-Go test, a modified Get-Up-and-Go test, staff's global judgement and fall history in evaluating fall risk in residential care facilities. *Age and Ageing*, 37, 442 - 448.
- Okochi, J., Utsunomiya, S., & Takahashi, T. (2005). Health measurement using the ICF: test-retest reliability study of ICF codes and qualifiers in geriatric care. *Health and Quality of Life Outcomes*, 3, 1 - 13.
- Olivier, N., & Rockmann, U. (2003). *Grundlagen der Bewegungswissenschaft und -lehre*. Schorndorf: Hofmann.
- Osholt-Corsten, M., & Schuntermann, M. (2011). Die ICD-10. In Deutsche Rentenversicherung Bund (Hrsg.), *Sozialmedizinische Begutachtung für die gesetzliche Rentenversicherung* (S. 66 - 68). Berlin: Springer.
- Palm, F., Urbanek, C., Rose, S., Buggle, F., Bode, B., Hennerici, M. G., ... & Piplack, K. O. (2010). Stroke incidence and survival in Ludwigshafen am Rhein, Germany: the Ludwigshafen stroke study (LuSSt). *Stroke*, 41, 1865 - 1870.
- Panzer, V. P., Wakefield, D. B., Hall, C. B., & Wolfson, L. I. (2011). Mobility assessment: sensitivity and specificity of measurement sets in older adults. *Archives of Physical Medicine and Rehabilitation*, 92, 905 - 912.
- Perell, K. L., Nelson, A., Goldman, R. L., Luther, S. L., Prieto-Lewis, N., Rubenstein, L. Z. (2001). Fall risk assessment measures: an analytic review. *The Journal of Gerontology. Series A, Biological Sciences and Medical Sciences*, 56, M761 - 766.

- Perkins, N. J., & Schisterman, E. F. (2006). The inconsistency of „optimal“ cut-points using two ROC based criteria. *American Journal of Epidemiology*, 163, 670 - 675.
- Pertille, A., Macedo, A. B., Dibai Filho, A. V., Rêgo, E. M., de Figueiredo Arrais, L. D., Negri, J. R., & Teodori, R. M. (2012). Immediate effects of bilateral grade III mobilization of the talocrural joint on the balance of elderly women. *Journal of Manipulative and Physiological Therapeutics*, 35, 549 - 555.
- Pfeifer, K., Sudeck, G., Geidl, W., & Tallner, A. (2013). Bewegungsförderung und Sport in der Neurologie – Kompetenzorientierung und Nachhaltigkeit. *Neurologie & Rehabilitation*, 19, 7 - 19.
- Pfeifer, K., & Sudeck, G. (2016). Körperliche Aktivität. In J. Bengel, & O. Mittag (Hrsg.), *Psychologie in der medizinischen Rehabilitation* (S. 215 - 228). Heidelberg: Springer.
- Pfingsten, M. (2004). Angstvermeidungs-Überzeugungen bei Rückenschmerzen. Gütekriterien und prognostische Relevanz des FABQ. *Der Schmerz*, 18, 17 - 27.
- Pfingsten, M. (2005). Bio-psycho-soziale Einflussfaktoren bei Rückenschmerz und Konsequenzen für die Bewegungstherapie. *Bewegungstherapie und Gesundheitssport*, 21, 152 - 158.
- Pisano, E. D., Hendrick, R. E., Yaffe, M. J., Baum, J. K., Acharyya, S., Cormack, J. B., ... & D'Orsi, C. J. (2008). Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology*, 246, 376 - 383.
- Podsiadlo, D., & Richardson, S. (1991). The Timed Up and Go – a test of basic functional mobility for frail elderly persons. *Journal of American Geriatric Society*, 39, 142 - 148.
- Pondal, M., & del Ser, T. (2008). Normative data and determinants for the timed “Up and Go” test in a population-based sample of elderly individuals without gait disturbances. *Journal of Geriatric Physical Therapy*, 31, 57 - 63.
- Pospeschill, M. (2010). *Testtheorie, Testkonstruktion, Testevaluation*. München, Basel: Reinhardt.
- Rantanen, T., Era, P., & Heikkinen, E. (1994). Maximal isometric strength and mobility among 75-year-old men and women. *Age and Ageing*, 23, 132 - 137.
- Rantanen, T., Guralnik, J. M., Foley, D., Masaki, K., Leveille, S., Curb, J. D., & White, L. (1999). Midlife hand grip strength as a predictor

- of old age disability. *Journal of the American Medical Association*, 281, 558 - 560.
- Rantanen, T. (2003). Muscle strength, disability and mortality. *Scandinavian Journal of Medicine & Science in Sports*, 13, 3 - 8.
- Reisberg, B., Cerris, S. H., De León, M. J., & Crook, T. (1982). The global deterioration scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry*, 139, 1136 - 1139.
- Regterschot, G. R. H., Folkersma, M., Zhang, W., Baldus, H., Stevens, M., & Zijlstra, W. (2014). Sensitivity of sensor-based sit-to-stand peak power to the effects of training leg strength, leg power and balance in older adults. *Gait & Posture*, 39, 303 - 307.
- Reitsma, J. B., Rutjes, A. W. S., Khan, K. S., Coomarasamy, A., & Bossuyt, P. M. (2009). A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *Journal of Clinical Epidemiology*, 62, 797 - 806.
- Ries, J. D., Echternach, J. L., Nof, L., & Gagnon Blodgett, M. (2009). Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Physical Therapy*, 89, 569 - 579.
- Robert Koch-Institut (RKI) (2013). *Arthrose. Gesundheitsberichterstattung des Bundes*. Berlin: RKI.
- Robert Koch-Institut (RKI) (2015). *Gesundheit in Deutschland. Gesundheitsberichterstattung des Bundes. Gemeinsam getragen von RKI und Destatis*. Berlin: RKI.
- Roberts, H. C., Denison, H. J., Martin, H. J., Patel, H. P., Syddall, H., Cooper, C., & Sayer, A. A. (2011). A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age and Ageing*, 40, 423 - 429.
- Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine*, 45, 1273 - 1284.
- Rockwood, K., Awalt, E., Carver, D., & MacKnight, C. (2000). Feasibility and measurement properties of the Functional Reach and the Timed Up and Go Test in the Canadian Study of Health and Aging. *The Journal of Gerontology Series A Biological Sciences and Medical Sciences*, 55, M70 - M73.
- Rose, D. J., Jones, C. J., & Lucchese, N. (2002). Predicting the probability of falls in community-residing older adults using the 8-foot up-

- and-go: a new measure of functional mobility. *Journal of Aging and Physical Activity*, 10, 466 - 475.
- Roth, K. (1999). Die fähigkeitsorientierte Betrachtungsweise. In K. Roth, & K. Willimchik (Hrsg.), *Bewegungswissenschaft* (S. 227 - 288). Reinbek/Hamburg: Rowohlt.
- Roth, K., & Roth, C. (2009). Entwicklung motorischer Fertigkeiten. In J. Baur, K. Bös, A. Conzelmann, & R. Singer (Hrsg.), *Handbuch Motorischer Entwicklung* (2. Aufl.) (S. 227 - 247). Schorndorf: Hofmann.
- Roth, K., & Willimczik, K. (1999). *Bewegungswissenschaft*. Reinbek/Hamburg: Rowohlt.
- Rubim, V. S. M., Drumond Neto, C., Romeo, J. L. M., & Montera, M. W. (2006). Prognostic value of the six-minute walk test in heart failure. *Arquivos Brasileiros de Cardiologia*, 86, 120 - 125.
- Sackett, D. L., & Haynes, R. B. (2002). The architecture of diagnostic research. *British Medical Journal*, 324, 539 - 541.
- Sakurai, R., Fujiwara, Y., Saito, K., Fukaya, T., Kim, M. J., Yasunaga, M., ... & Muraki, E. (2013). Effects of a comprehensive intervention program, including hot bathing, on overweight adults: a randomized controlled trial. *Geriatrics & Gerontology International*, 13, 638 - 645.
- Salbach, N. M., Mayo, N. E., Wood-Dauphinee, S., Hanley, J. A., Richards, C. L., & Côte, R. (2004). A task-oriented intervention enhances walking distance and speed in the first year post stroke: a randomized controlled trial. *Clinical Rehabilitation*, 18, 509 - 519.
- Sallinen, J., Stenholm, S., Rantanen, T., Heliövaara, M., Sainio, P., & Koskinen, S. (2010). Hand-grip strength cut points to screen older persons at risk for mobility limitation. *Journal of the American Geriatrics Society*, 58, 1721 - 1726.
- Samson, M. M., Meeuwssen, I. B., Crowe, A., Dessens, J. A., Duursma, S. A., & Verhaar, H. J. (2000). Relationships between physical performance measures, age, height and body weight in healthy adults. *Age and Ageing*, 29, 235 - 242.
- Samson, M. M., Crowe, A., De Vreede, P. L., Dessens, J. A. G., Duursma, S. A., & Verhaar, H. J. J. (2001). Differences in gait parameters at a preferred walking speed in healthy subjects due to age, height and body weight. *Aging Clinical and Experimental Research*, 13, 16 - 21.
- Sasaki, H., Kasagi, F., Yamada, M., & Fujita, S. (2007). Grip strength predicts cause-specific mortality in middle-aged and elderly persons. *The American Journal of Medicine*, 120, 337 - 342.

- Savva, G. M., Donoghue, O. A., Horgan, F., O'Regan, C., Cronin, H., & Kenny, R. A. (2012). Using timed up-and-go to identify frail members of the older population. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 68, 441 - 446.
- Scale, B. B., & Talking, S. W. W. (2006). How to identify potential fallers in a stroke unit: validity indexes of four test methods. *Journal of Rehabilitation Medicine*, 38, 186 - 191.
- Schaaf, H. (2017). *Morbus Menière: Schwindel-Hörverlust-Tinnitus: eine psychosomatisch orientierte Darstellung*. Heidelberg: Springer.
- Schacht, E., & Ringe, J. D. (2012). Alfacalcidol improves muscle power, muscle function and balance in elderly patients with reduced bone mass. *Rheumatology International*, 32, 207 - 215.
- Scharhag-Rosenberger, F., Becker, T., Streckmann, F., Schmidt, K., Berling, A., Bernardi, A., ... & Steindorf, K. (2014). Studien zu körperlichem Training bei onkologischen Patienten: Empfehlungen zu den Erhebungsverfahren. *Deutsche Zeitschrift für Sportmedizin*, 65, 304 - 313.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8, 350.
- Schoene, D., Wu, S. M. S., Mikolaizak, A. S., Menant, J. C., Smith, S. T., Delbaere, K., & Lord, S. R. (2013). Discriminative ability and predictive validity of the timed Up and Go test in identifying older people who fall: systematic review and meta-analysis. *Journal of the American Geriatrics Society*, 61, 202 - 208.
- Schoppen, T., Boonstra, A., Groothoff, J. W., de Vries, J., Göeken, L. N., & Eisma, W. H. (1999). The Timed "up and go" test: reliability and validity in persons with unilateral lower limb amputation. *Archives of Physical Medicine and Rehabilitation*, 80, 825 - 828.
- Schuck, P. (2000). Designs und Kennziffern zur Ermittlung der Änderungssensitivität von Fragebogen in der gesundheitsbezogenen Lebensqualitätsforschung. *Zeitschrift für medizinische Psychologie*, 9, 125 - 130.
- Schuntermann, M. F. (2005). Die Internationale Klassifikation für Funktionsfähigkeit, Behinderung und Gesundheit (ICF) der Weltgesundheitsorganisation (WHO). Kurzeinführung. *Physikalische Medizin, Rehabilitationswissenschaften, Kurortmedizin*, 11, 229 - 230.
- Shimada, H., Suzukawa, M., Tiedemann, A., Kobayashi, K., Yoshida, H., & Suzuki, T. (2009). Which neuromuscular or cognitive test is the optimal screening tool to predict falls in frail community-dwelling older people? *Gerontology*, 55, 532 - 538.



- Shumway-Cook, A., Brauer, S., & Woollacott, M. (2000). Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Physical Therapy*, 80, 896 - 903.
- Shumway-Cook, A., Ciol, M. A., Yorkston, K. M., Hoffman, J. M., & Chan, L. (2005). Mobility limitations in the Medicare population: prevalence and sociodemographic and clinical correlates. *Journal of the American Geriatrics Society*, 53, 1217 - 1221.
- Shumway-Cook, A., & Woollacott, M. (2007). *Motor control. Translating research into clinical practice*. Philadelphia: Lippencott Williams & Wilkins.
- Sikjaer, T., Rolighed, L., Hess, A., Fuglsang-Frederiksen, A., Mosekilde, L., & Rejnmark, L. (2014). Effects of PTH (1 - 84) therapy on muscle function and quality of life in hypoparathyroidism: results from a randomized controlled trial. *Osteoporosis International*, 25, 1717 - 1726.
- Simonsick, E. M., Newman, A. B., Visser, M., Goodpaster, B., Kritchevsky, S. B., Rubin, S., ... & Health, Aging and Body Composition Study. (2008). Mobility limitation in self-described well-functioning older adults: importance of endurance walk testing. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 63, 841 - 847.
- Smith, J., Fleeson, W., Geiselmann, B., Settersten, R., & Kunzmann, U. (1996). Wohlbefinden im hohen Alter: Vorhersagen aufgrund objektiver Lebensbedingungen und subjektiver Bewertung. In K. U. Mayer, & P. B. Baltes (Hrsg.), *Die Berliner Altersstudie. Ein Projekt der Berlin-Brandenburgischen Akademie der Wissenschaften* (S. 487 - 523). Berlin: Akademie-Verlag.
- Soety, N. R., Faruquee, M. H., Lahiry, S., Chaklader, M. A., & Yasmin, N. (2014). Handgrip strength and nutritional status of female garment workers in Dhaka City. *International Public Health Forum*, 14, 18 - 22.
- Sosa, E. S., Groeneveld, I. F., Gonzalez-Saiz, L., Lopez-Mojares, L. M., Villa-Asensi, J. R., Gonzalez, M. I. B., ... & Lucia, A. (2012). Intrahospital weight and aerobic training in children with cystic fibrosis: a randomized controlled trial. *Medicine & Science in Sports & Exercise*, 44, 2 - 11.
- Sosnoff, J. J., Socie, M. J., Boes, M. K., Sandroff, B. M., Pula, J. H., Suh, Y., ... & Motl, R. W. (2011). Mobility, balance and falls in persons with multiple sclerosis. *PloS one*, 6, 1 - 5.
- Southard, V., Dave, A., & Douris, P. (2010). Exploring the Role of Body Mass Index on Balance Reactions and Gait in Overweight Seden-

- tary Middle-aged Adults: A Pilot Study. *Journal of Primary Care & Community Health*, 1, 178 - 83.
- Statistisches Bundesamt (2015). *Bevölkerung Deutschlands bis 2060. Ergebnisse der 13. koordinierten Bevölkerungsvorausberechnung*. Wiesbaden: Destatis.
- Steffen, T. M., Hacker, T. A., & Mollinger, L. (2002). Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Physical Therapy*, 82, 128 - 137.
- Stuck, A. E., Walthert, J. M., Nikolaus, T., Büla, C. J., Hohmann, C., & Beck, J. C. (1999). Risk factors for functional status decline in community-living elderly people: a systematic literature review. *Social Science & Medicine*, 48, 445 - 469.
- Stucki, G., Cieza, A., Ewert, T., Kostanjsek, N., Chatterji, S., & Üstün, T. B. (2002). Application of the International Classification of Functioning, Disability and Health (ICF) in clinical practice. *Disability and Rehabilitation*, 24, 281 - 282.
- Stucki, G., Cieza, A., & Melvin, J. (2007). The international classification of functioning, disability and health: A unifying model for the conceptual description of the rehabilitation strategy. *Journal of Rehabilitation Medicine*, 39, 279 - 285.
- Such, U., & Meyer, T. (2010). Die maximale Herzfrequenz. *Deutsche Zeitschrift für Sportmedizin*, 61, 310 - 311.
- Sucuoglu, H., Tuzun, S., Akbaba, Y. A., Uludag, M., & Gokpinar, H. H. (2015). Effect of whole-body vibration on balance using posturography and balance tests in postmenopausal women. *American Journal of Physical Medicine & Rehabilitation*, 94, 499 - 507.
- Tack, W.-H. (1976). Diagnostik als Entscheidungshilfe. In K. Pawlik (Hrsg.), *Diagnose der Diagnostik* (S. 103 - 129). Stuttgart: Klett.
- Thompson, P. D., Arena, R., Riebe, D., & Pescatello, L. S. (2013). ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription. *Current Sports Medicine Reports*, 12, 215 - 217.
- Tinetti, M. E., Williams, T. F., Mayewski, R. (1986). Fall risk index for elderly patients based on number of chronic disabilities. *The American Journal of Medicine*, 80, 429 - 434.
- Tinetti, M. E., Richman, D., & Powell, L. (1990). Falls efficacy as a measure of fear of falling. *Journal of Gerontology*, 45, 239 - 243.

- Thrane, G., Joakimsen, R. M., & Thornquist, E. (2007). The association between timed up and go test and history of falls: the Tromsø study. *BMC Geriatrics*, 7, 1 - 7.
- Tsuji, T., Kitano, N., Tsunoda, K., Himori, E., Okura, T., & Tanaka, K. (2014). Short-term effects of whole-body vibration on functional mobility and flexibility in healthy, older adults: a randomized cross-over study. *Journal of Geriatric Physical Therapy*, 37, 58 - 64.
- Vance, C. G. T., Rakel, B. A., Blodgett, N. P., DeSantana, J. M., Amendola, A., Zimmerman, M. B., ... & Sluka, K. A. (2012). Effects of transcutaneous electrical nerve stimulation on pain, pain sensitivity, and function in people with knee osteoarthritis: a randomized controlled trial. *Physical Therapy*, 92, 898 - 910.
- Verdijk, L. B., Snijders, T., Holloway, T. M., Van, J. K., & Van, L. L. (2016). Resistance training increases skeletal muscle capillarization in healthy older men. *Medicine and Science in Sports and Exercise*, 48, 2157 - 2164.
- Vereeck, L., Wuyts, F., Truijen, S., & Van de Heyning, P. (2008). Clinical assessment of balance: normative data, and gender and age effects. *International Journal of Audiology*, 47, 67 - 75.
- Visser, M., Goodpaster, B. H., Kritchevsky, S. B., Newman, A. B., Nevitt, M., Rubin, S. M., ... & Harris, T. B. (2005). Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60, 324 - 333.
- Warburton, D. E., Gledhill, N., & Quinney, A. (2001). Musculoskeletal fitness and health. *Canadian Journal of Applied Physiology*, 26, 217 - 237.
- Wevers, L., van de Port, I., Vermue, M., Mead, G., & Kwakkel, G. (2009). Effects of task-oriented circuit class training on walking competency after stroke: a systematic review. *Stroke*, 40, 2450 - 2459.
- Wendel, K., Ståhl, A., Risberg, J., Pessah-Rasmussen, H., & Iwarsson, S. (2010). Post-stroke functional limitations and changes in use of mode of transport. *Scandinavian Journal of Occupational Therapy*, 17, 162 - 174.
- Whiting, P., Rutjes, A. W., Reitsma, J. B., Glas, A. S., Bossuyt, P. M., & Kleijnen, J. (2004). Sources of variation and bias in studies of

- diagnostic accuracy: a systematic review. *Annals of Internal Medicine*, 3, 189 - 202.
- Whitney, S. L., Hudak, M. T., & Marchetti, G. F. (2000). The dynamic gait index relates to self-reported fall history in individuals with vestibular dysfunction. *Journal of Vestibular Research*, 10, 99 - 105.
- Whitney, S. L., Marchetti, G. F., Schade, A., & Wrisley, D. M. (2004). The sensitivity and specificity of the Timed "Up & Go" and the Dynamic Gait Index for self-reported falls in persons with vestibular disorders. *Journal of Vestibular Research*, 14, 397 - 409.
- Whitney, J. C., Lord, S. R., & Close, J. C. (2005). Streamlining assessment and intervention in a falls clinic using the Timed Up and Go Test and Physiological Profile Assessments. *Age and Ageing*, 34, 567 - 571.
- Williams, E. N., Carroll, S. G., Reddihough, D. S., Phillips, B. A., & Galea, M. P. (2005). Investigation of the timed "up & go" test in children. *Developmental Medicine and Child Neurology*, 47, 518 - 524.
- Williams, P. E., & Goldspink, G. (1978). Changes in sarcomere length and physiological properties in immobilized muscle. *Journal of Anatomy*, 127, 459 - 468.
- Willimczik, K., & Singer, R. (2009). Motorische Entwicklung: Gegenstandsreich. In J. Baur, K. Bös, A. Conzelmann, & R. Singer (Hrsg.), *Handbuch Motorische Entwicklung* (S. 15 - 24). Schorndorf: Hofmann.
- Wilson, J. M. G., Jungner, G., & World Health Organization (WHO) (1968). Principles and practice of screening for disease. Geneva: WHO.
- World Health Organization (WHO) (2001). *International Classification of Functioning, Disability and Health: ICF*. Geneva: WHO.
- World Health Organization (WHO) (2003). *The burden of musculoskeletal conditions at the start of the new millennium. Report of a WHO Scientific Group Technical Report Series*. Geneva: WHO.
- World Health Organization (WHO) (2006). Constitution of the World Health Organization. Basic Documents. Zugriff unter [http://www.who.int/governance/eb/who\\_constitution\\_en.pdf](http://www.who.int/governance/eb/who_constitution_en.pdf)
- Wright, A. A., Cook, C. E., Baxter, G. D., Dockerty, J. D., & Abbott, J. H. (2011). A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. *Journal of Orthopaedic & Sports Physical Therapy*, 41, 319 - 327.
- Wydra, G. (1996). *Gesundheitsförderung durch sportliches Handeln*. Schorndorf: Hofmann.

- Wydra, G. (2006a). Assessmentverfahren. In G. Wydra, H. Winchenbach, M. Schwarz, & K. Pfeifer (Hrsg.), *Assessmentverfahren in Gesundheitssport und Bewegungstherapie. Messen, Testen, Beurteilen, Bewerten* (S. 9 - 21). Hamburg: Czwalina.
- Wydra, G. (2006b). Normierung der motorischen Leistungsfähigkeit. *Bewegungstherapie und Gesundheitssport*, 22, 223 - 227.
- Wydra, G., & Kaczmarek, C. (2016). Screeningtests in Gesundheits- und Rehasport. In H. Wäsche, G. Sudeck, R. S. Kähler, L. Vogt, & A. Woll (Hrsg.), *Bewegung, Raum und Gesundheit – Wechselwirkungen im Spannungsfeld geänderter Lebensbedingungen und Mobilitäten* (Abstractband, S. 84). Karlsruhe: dvs.
- Wydra, G., & Kaczmarek, C. (2017). Von der Fähigkeits- zur Mobilitätsorientierung in Gesundheits- und Rehasport. In H. Wäsche, G. Sudeck, R. S. Kähler, L. Vogt, & A. Woll (Hrsg.), *Bewegung, Raum und Gesundheit* (S. 111 - 119). Hamburg: Feldhaus.
- Yang, Y.-R., Wang, R.-Y., Lin, K.-H., Chu, M.-Y., & Chan, R.-C. (2006). Task-oriented progressive resistance strength training improves muscle strength and functional performance in individuals with stroke. *Clinical Rehabilitation*, 20, 860 - 870.
- Yeomans, K. A., & Golder, P. A. (1982). The Guttman-Kaiser Criterion as a predictor of the number of common factors. *The Statistician*, 31, 221 - 229.
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3, 32 - 35.
- Zampieri, C., Salarian, A., Carlson-Kuhta, P., Aminian, K., Nutt, J. G., & Horak, F. B. (2010). The instrumented timed up and go test: potential outcome measure for disease modifying therapies in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 81, 171 - 176.
- Zhou, X. H., Obuchowski, N. A., & McClish, D. K. (2002). *Statistical Methods in Diagnostic Medicine*. New York: Wiley.
- Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115, 654 - 657.